Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase

Article

Accepted Version

Topological impact of non-canonical DNA structures on Klenow Fragment of DNA polymerase

Shuntaro Takahashi[1], John A. Brazier[2], and Naoki Sugimoto[1,3*]

[1] Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Kobe 650-0047, Japan

[2] Department of Pharmacy, University of Reading, Whiteknights, Reading, Berkshire RG6 6AD, UK

[3] Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Kobe 650-0047, Japan

* To whom correspondence should be addressed. Tel: (+81)78-303-1457; Fax: (+81)78-303-1495; Email: sugimoto@konan-u.ac.jp

Keywords: *Replication, i-motif, G-quadruplex, Thermodynamics, Molecular crowding*

ABSTRACT

Non-canonical DNA structures which stall DNA replication can cause errors in genomic DNA. Here we investigated how the non-canonical structures formed by sequences in genes associated with a number of diseases impacted DNA polymerization by Klenow Fragment of DNA polymerase. Replication of DNA sequence forming i-motif from telomere, hypoxia-induced transcription factor and an insulin-linked polymorphic region (ILPR) was effectively inhibited. On the other hand, replication of mixed-type G-quadruplex from telomere was less inhibited than that of anti-parallel-type or parallel-type. Interestingly, the i-motif was a better inhibitor of replication than were mixed-type G-quadruplexes or hairpin structures, even though all had similar thermodynamic stabilities. These results indicate that both stability and topology of structures formed in DNA templates impact the processivity of a DNA polymerase. This suggests that i-motif formation may trigger genomic instability by stalling replication of DNA causing intractable diseases.

SIGNIFICANCE STATEMENT

Alterations in non-duplex structures could play roles during DNA replication in the progression of cancer and other intractable diseases. These non-canonical structures differ topologically from each other. However, the role of these differences in diseases remains unknown. In this study, we found that the presence of i-motif structures in the template caused the DNA polymerase to stall immediately before these structures. The i-motif structures are more efficient than other structures such as G-quadruplexes and hairpins, although their thermodynamic stabilities are similar. This inhibition effect of the DNA polymerase was regulated by molecular crowding, which mimics conditions in the cell. Therefore, it is possible that the i-motif could impede DNA replication or repair and cause genomic instability.

/body

INTRODUCTION

Non-canonical intramolecular structures of nucleic acids, such as a triplex and a quadruplex, are stabilized under conditions that mimic the crowded cellular conditions,(1) and have been detected in cells.(2, 3) *In vitro* and *in vivo*, guanine-quadruplex (G-quadruplex or G4) formation

inhibits transcription and translation of template nucleic acids.(4-8) It is possible that the non-canonical structures act as “functional codes”, triggered by different molecular environments, which regulate gene expression epigenetically.(6, 7) As sequences capable of forming the non-canonical structures are found in telomeres and promoter regions of known oncogenes, alterations of the non-canonical structures could play important roles in the progression of cancer and in other diseases.(9)

One of the remarkable features of non-canonical structures of nucleic acids is the diversity of topologies. In the case of G4s, antiparallel, mixed, and parallel type structures have been characterized (Figs. 1A-C). The sequences complementary to regions capable of G4 formation are composed of tandem repeats of cytosine, and these C-rich regions can form a different type of tetraplex topology, which is the i-motif.(10, 11) An intramolecular i-motif is formed upon the interaction of four C-rich regions. The structure has three loops and two parallel-stranded hairpin-like units stabilized by $C:C^+$ base pairs that are vertically intercalated antiparallel to one another (Figs. 1D,E). The i-motif structures are categorized into Class I and Class II based on the lengths of the loops. C-rich sequences are found in telomeres and in the promoter regions of about 40% of human genes.(12, 13) As the i-motif is stabilized by hydrogen bonding between cytosine and protonated cytosine ($C:C^+$),(14, 15) acidic conditions stabilize the structure. As the i-motif can mediate transcriptional regulation of B-cell lymphoma 2 (Bcl2) oncogene in cells,(16) the environment inside these cells might be favorable to i-motif formation. Interestingly, the promoter region of the gene that encodes hypoxia-induced transcription factor, Hif1a, which is highly expressed in cancer cells, forms a very stable i-motif structure *in vitro*.(12, 17) The thermodynamic stability of the i-motif structure with Hif1a sequence under slightly acidic conditions is comparable to that of G4 formed by the opposite strand.(18) Since G4 formation in the template DNA causes breakage of the genomic DNA due to stalling of DNA polymerase during replication,(19, 20) the presence of an i-motif on the template DNA may be responsible for the mechanism of intractable diseases including cancer. Recently, the stability of G4 *in vitro* did not always depend on the frequency of genomic instability, which may indicate that different topologies affect replication. However, the effects of these topological differences, particularly the i-motif, remain unknown.

In this study, we investigated the effect of structures formed in the template strand on the replication reaction. Under slightly acidic conditions, i-motif formation decreased the rate of replication by Klenow fragment DNA polymerase (KF). Gel electrophoretic analyses revealed

that the DNA polymerase stalled immediately before the i-motif forming region, indicating that the i-motif is an obstruction to the polymerase. We compared the inhibition resulting from i-motif formation to that resulting from G4 and hairpin formation and found that the level of inhibition was determined not only by the stability but also by the topology of non-canonical DNA structures. The i-motif, anti-parallel, and parallel G4s repressed replication more effectively than mixed G4 or hairpins with similar thermodynamic stabilities. These findings suggest that not only G4 but also the i-motif can induce genomic instability more than other non-canonical structures, which causes intractable diseases.

RESULTS

DNA replication reaction. In the template DNA the structure-forming sequence was adjacent to the region complementary to the primer (SI Appendix, Fig. S1). The 5' terminus of primer DNA was labeled with fluorescein (FAM) in order to enable quantification of product formation. The gap between the region of a non-canonical structure and the primer binding region was four bases, which is a length selected to ensure that the structure adopted does not interfere with the formation of the elongation complex of KF and initial polymerization.(21) This design enabled us to discriminate stalled product from unreacted primer and full-length product. The i-motif forming sequences used are derived from the human telomeric sequence $(C_3TA_2)_4$, the promoter region of Hif1a gene, a region complementary to a portion of the insulin-linked polymorphic region (cILPR), which is a regulatory sequence upstream of the gene encoding insulin, and the complementary sequence of an oncogene of Bcl2 (cBcl2). Also tested were G4-forming sequences $(T_2AG_3)_4$, Q6, and ILPR (complementary G-rich sequence of cILPR), and hairpin-forming sequences (H1, H2, and H3). Q6 possesses four repeats of the G_4 unit linked with the loop of the thrombin binding aptamer.(6) The control was a linear sequence not expected to form any stable structures. All the sequences are shown in SI Appendix, Table S1. Formation of the i-motif and G4 structures which have different topologies were confirmed by circular dichroism (CD) analyses at 37 °C in the buffer used for replication assays (SI Appendix, Fig. S2, S3).

Replication of DNA containing an i-motif structure.

Although inhibition of DNA replication by a G4 formed within a template DNA strand has been studied before,(22) there have been no reports on the effects of i-motif structures on DNA replication. Therefore, we first examined replication of a template that contained the human telomeric sequence $(C_3TA_2)_4$ forming the i-motif structure. The replication reaction was carried out with 1 μM KF and 250 μM dNTPs for 0.5 min at 37 °C, and the progress of the reaction was analyzed by quantification of fluorescently labeled products on denaturing polyacrylamide gel electrophoresis (PAGE). At pH 7.0, a condition in which $(C_3TA_2)_4$ does not form the i-motif structure, the fluorescently labeled primer was completely converted to a longer product within 0.5 min (Fig. 2A). By staining of DNA using SYBR Gold dye, we confirmed that the fluorescently labeled replication product had the same mobility as the template DNA (SI Appendix, Fig. S4). Thus, the replication of the $(C_3TA_2)_4$-containing template yielded a full-length product. Under slightly acidic condition at pH 6.0, replication of the i-motif-forming template was repressed (Fig. 2B). A band that migrated just above the primer band was observed in the reaction with the $(C_3TA_2)_4$-containing template at pH 6.0 (Fig. 2B). The presence of this band indicates that the replication was stalled after primer extension of less than 10 nucleotides (SI Appendix, Fig. S4). Such a short band was not observed in the case of the replication of the Linear template (Fig. 2C). These data suggest that the i-motif structure formed at low pH transiently stalled replication immediately before the structure. When the experiment was performed on a template containing the i-motif forming sequence found in the promoter region of Hif1a gene, which forms a particularly stable i-motif structure *in vitro*(12, 17), the results were more striking. At pH 7.0, the replication of the Hif1a template was efficient (Fig. 2D), but at pH 6.0 virtually no full-length product was detected (Fig. 2E). We did observe a small amount of a shorter product, indicating that the replication reaction was inhibited after elongation of less than 10 nucleotides (SI Appendix, Fig. S4). It was shown previously that KF stalls immediately before a G4 structure on a template.(22) To confirm that the shorter products resulted from stalling of the polymerase during progressive replication, we analyzed the replication from FAM-labeled primer S, which was 10 bases shorter than the original primer. Reactions using this primer generated the same length stalled product as those using the longer primer, indicating that the non-canonical structure did not disrupt template loading (SI Appendix, Fig. S5). Thus, during replication of the Hif1a template DNA, we confirmed that KF was stalled by the i-motif structure formed within the template. We also confirmed replication stall by T7 DNA polymerase which replicates genome of T7 phage (SI Appendix, Fig. S6). Therefore, stalling of replication at an i-motif can be considered to generally occur.

Dependence of replication efficiency on i-motif stability.

To understand the relationship between the replication efficiency and i-motif stability, a UV melting analysis was performed on strands containing only the i-motif-forming region of the template sequences (Table 1, SI Appendix, Table S2 and Figs. S7, S8). At pH 6.0, $(C_3TA_2)_4$ had a melting temperature (T_m) of 31.4 °C and a free energy change at 37 °C ($-\Delta G°_{37}$) of -0.70 kcal mol^{-1}, whereas the Hif1a sequence had a T_m of 57.3 °C and a $-\Delta G°_{37}$ of 3.1 kcal mol^{-1} (Table 1). These data suggest that the stalling effect of KF is related to the magnitude of $-\Delta G°_{37}$. We previously showed that transcription by RNA polymerase was halted at a G4 structure in the template DNA(6) and that the magnitude of the inhibition depended on the stability of G4. There has been no quantitative analysis of the effect of stabilities of non-canonical structures present in the template DNA on replication, but a previous study showed that replication was inhibited at a G4 and that the inhibition was more efficient in the presence of a G4 binding ligand expected to increase the stability of the G4 structure.(5) The data reported here suggest that i-motif can also inhibit replication in a stability-dependent manner.

Effect of molecular crowding on the replication of i-motif-forming DNA.

To further investigate the replication of i-motif-forming DNA in cell-mimicking conditions, we examined replication in more physiologically relevant conditions in the presence of salt and crowding reagents using polyethylene glycol (PEG). In the presence of 100 mM KCl and 20 wt% PEG200 (average molecular weight is 200) at pH 6.0, the amount of full-length product replicated from the template containing the $(C_3TA_2)_4$ sequence was quite similar to that in the absence of PEG200 (SI Appendix, Fig. S9A). Since the higher molecular weight of PEG shows more stabilization of i-motif due to the excluded volume effect, (23, 24) the replication might not be effectively stalled by the addition of 20 wt% PEG200. In fact, the addition of 20 wt% PEG200 did not increase thermal stability of $(C_3TA_2)_4$ i-motif, whose T_m was 28.7°C and $-\Delta G°_{37}$ was -1.4 kcal mol^{-1} (SI Appendix, Table S2). In the presence of 20 wt% PEG1000 (average molecular weight is 1000) at pH 6.0, a short product was detected on PAGE (SI Appendix, Fig. S9B), indicating that KF was stalled immediately before the i-motif structure. The stabilizing effect of PEG1000 was observed in the melting temperature of $(C_3TA_2)_4$ (SI Appendix, Fig. S8 and Table

S2). The T_m in the presence of 20 wt% PEG1000 was 37.3 °C and the $-\Delta G°_{37}$ was 0.1 kcal mol^{-1}. In the case of Hif1a sequence, the thermodynamic stability was also not affected in the presence of PEG200 (T_m = 51.2 °C and $-\Delta G°_{37}$ = 3.1 kcal mol^{-1}) but increased in the presence of PEG1000 (T_m = 59.7 °C and $-\Delta G°_{37}$ = 4.2 kcal mol^{-1}) (SI Appendix, Table S2 and Fig. S8). The Hif1a sequence stalled KF immediately before the i-motif structure in the presence of PEG1000, although PEG200 did not have an effect (SI Appendix, Figs. S9C,D). Interestingly, stalled replication could be observed even at pH 7.0 in the presence of 100 mM KCl and 20 wt% PEG1000 (SI Appendix, Fig. S9E). Although the detailed mechanism of stabilization of the i-motif structure by PEG1000 is unknown, the higher molecular weight PEG shifts the pK_a of N3 in the cytosine base.(23, 24) As the i-motif structure is very compact compared to a random coil, a more pronounced excluded volume effect of higher molecular weight PEG1000 compared to that of PEG200 may contribute to the observed effect of PEG1000 on the replication of an i-motif-forming DNA. In summary, our results suggest that the cellular condition crowded with biomacromolecules stabilizes the i-motif structure, which facilitates the inhibition of replication.

Replication rate is inversely correlated with stability of i-motif.

To further analyze the correlation between the thermodynamic stability and replication rate, we evaluated the amount of product as a function of time (SI Appendix, Fig. S10). In the reaction with the $(C_3TA_2)_4$ template in the absence of PEG at pH 6.0, stalled product was observed at 0.16 min (SI Appendix, Fig. S10A). The amount of stalled product was decreased at 0.33 min and had almost disappeared by 1 min. On the other hand, full-length product was detected at 0.33 min, and the conversion to full-length product was complete at 2 min. In the case of the replication of Hif1a sequence, stalled product was detected at 0.5 min, and the amount gradually decreased with time. A small amount of full-length product was detected at 1 min and the amount increased at longer time points (SI Appendix, Fig. S10B). For both $(C_3TA_2)_4$ and Hif1a templates, the increase in full-length product corresponded to a decrease in amount of the stalled product (Fig. 3A). The generation curves of full-length product were analyzed by global fitting (SI Appendix, Fig. S11) and rate constants, k_s (min^{-1}) for the reaction step of the stall of replication including the prior replication to the stalling position and k_f (min^{-1}) for the reaction step of generating full-length product after resolving the replication stall were calculated. The k_s values were 4.5 min^{-1} and 0.38 min^{-1} at 37 °C and pH 6.0 for the templates containing the

$(C_3TA_2)_4$ sequence and the Hif1a sequence, respectively (Table 1). These results indicate that KF took about 12 times longer time to overcome the more stable i-motif formed by the Hif1a sequence than to overcome that formed by the $(C_3TA_2)_4$ sequence. The k_f values were 4.5 min^{-1} and 6.3 min^{-1} at 37 °C and pH 6.0 for the templates containing the $(C_3TA_2)_4$ sequence and the Hif1a sequence, respectively (SI Appendix, Table S2), suggesting that the replication stall did not occur for the $(C_3TA_2)_4$ sequence but the stalling step was the rate-determining step in the replication of Hif1a sequence.

Fig. 3B shows a plot of the logarithm of replication rate constant ($\ln k_s$) versus $-\Delta G°_{37}$ for the template sequences investigated. The $-\Delta G°_{37}$ value for the Linear control template was assumed to be 0. This plot includes data on an additional i-motif formed by cILPR. cILPR is a portion of the complementary sequence of insulin-linked polymorphic region, which forms a previously characterized i-motif.(18) The values of $\ln k_s$ decreased with increasing $-\Delta G°_{37}$, and there is a good linear correlation. As $\Delta G°$ is equal to $-RT \ln K$ (R is a gas constant and K is the equilibrium constant), $\Delta G°$ also can be described as $-RT \ln(k_1/k_{-1})$, where k_1 is the rate constant of structure formation and k_{-1} is the deformation rate constant. Because the stabilities of DNA structures depend on hydrogen bonding and stacking interactions between bases, k_1 values for DNAs of similar chain lengths are similar, and thus k_{-1} is a dominant factor in stability.(25) Therefore, the linear relationship between $\ln k_s$ and $-\Delta G°_{37}$ indicates that the replication rate through an i-motif is mainly proportional to the rate of the unfolding of the i-motif structure (k_{-1}). As the temperature was constant in our analysis, we infer that $\ln k_s$ is proportional to the activation free energy $\Delta G‡$. The $\Delta G‡$ of the unfolding of a DNA duplex ($\Delta G‡_{off}$) has a linear correlation with the stability of the DNA duplex ($-\Delta G°$),(26) and the activation free energy of dissociation of a DNA duplex depends on the sequence (GC content) and length.(27, 28) Therefore, to unwind structured DNAs, the polymerase must decrease the $\Delta G‡_{off}$ or rely on spontaneous dissociation of the structured DNA. It is possible that reduced enzyme processivity due to low pH and PEG may be the factor in the replication stall. Low pH and PEG do affect enzyme processivity because the k_s of Linear template at pH 7.0 in the absence of salt and PEG was 18 min^{-1}, which was about 5-fold larger than that at pH 6.0 (SI Appendix, Figs. S10, S11, and Table S2). However, the slope of the $\ln k_s$ vs $-\Delta G°_{37}$ plot of i-motifs, which indicates the magnitude of $\Delta G‡$ for the unwinding process of i-motif by KF, showed similar values at pH 6.0 in the absence of salt and PEG (-0.83), in the presence of 100 mM KCl without PEG (-0.73), and pH 7.0 in the presence of 100 mM KCl and 20 wt% PEG1000 (-0.55) (SI Appendix, Figs. S12-S14) (discussion of PEG200 in the next section). The

similar value obtained for the slope indicated that the unwinding of i-motif was dominated by the same mechanism, not by enzyme properties, in various conditions. Thus, the increased stalling at the i-motif was due to folding of the i-motif and not reduced enzyme processivity at low pH or in PEG.

Effect of topology of structured DNAs on the replication reaction.

We next analyzed the replication of template DNAs having sequences that can form different structural topologies. First, we investigated the replication of a template DNA containing a hairpin structure. The sequences contained stems of four (H1), nine (H2), and twelve (H3) base pairs, and all had a four-nucleotide loop characterized previously.(6) The $-\Delta G°_{37}$ values obtained from UV melting experiments in 40 mM MES (pH 6.0) and 8 mM $MgCl_2$ were 2.2, 4.0, and 8.3 kcal mol^{-1} for H1, H2, and H3, respectively (Table 1, SI Appendix, Table S2, and Fig. S7). Thus, each of these hairpins is more stable than the i-motif structure adopted by $(C_3TA_2)_4$, which has a $-\Delta G°_{37}$ of -0.70 kcal mol^{-1}. Interestingly, KF did not stall on any of the hairpin-containing templates (SI Appendix, Fig. S10).

We also analyzed replication of templates that contained regions able to form G4 in different KCl concentrations. For the template containing $(T_2AG_3)_4$ in the presence of 30 mM KCl at 37 °C and pH 6.0, stalled product was observed at 0.16 min (Figs. 3C, and SI Appendix, S10l) and the full-length product increased correspondingly. UV melting showed that the $-\Delta G°_{37}$ value of $(T_2AG_3)_4$ was 3.0 kcal mol^{-1} in the presence of 30 mM KCl (Table 1 and SI Appendix, Fig. S7). We also confirmed that the stalled product was due to a replication block at the G4 structure (SI Appendix, Fig. S5D). Interestingly, although Hif1a, H2, and $(T_2AG_3)_4$ in 30 mM KCl showed similar stabilities (Table 1), the k_s values differed, indicating that the replication stall depended on not only the stability, but also topology.

The plot of $\ln k_s$ versus $-\Delta G°_{37}$ for hairpin DNAs and $(T_2AG_3)_4$ at 1, 10, 30, and 50 mM KCl were linear as shown in the case of i-motif templates (Fig. 3B). The slope of the $\ln k_s$ versus $-\Delta G°_{37}$ is -0.048 for hairpin structures and was -0.28 for $(T_2AG_3)_4$, which is different from that of -0.83 for the i-motif, indicating that the activation free energy ($\Delta G°_{37}{}^{\ddagger}$) required to unwind the i-motif structures was about 17-fold or 3-fold higher than that required to unwind hairpin or $(T_2AG_3)_4$ G4, which forms a mixed structure (SI Appendix, Fig. S2). The data for the template containing the

G4 ILPR which forms parallel structure(29) and Q6 sequences at 1 mM KCl which forms anti-parallel structure also showed a different stability tendency and k_s value (Table 1). Those plots did not fall on the line with data for $(T_2AG_3)_4$-containing templates. This suggests that the topology of G4 affects the efficiency of unwinding. Interestingly, the data for the ILPR and Q6 template showed a correlation with data on i-motifs, suggesting that the i-motif is a strong block for replication as well as parallel/anti-parallel G4s.

To investigate the effect of i-motif topology on replication, we tested the Class II i-motif cBcl2 sequence (Fig. 1E) that is found in the promoter region of the Bcl2 oncogene. In this case, we observed two major bands shorter than the full-length product (SI Appendix, Fig. S15A). The Bcl2 sequence can adopt i-motif structures of different conformations and the sequence can also form a hairpin structure, due to the presence of six cytosine tracts. If multiple conformers are present more than one stalled product might be observed. Moreover, there were unexpected smeared bands above the full-length product band. Since a nascent DNA product from the replication of triplet repeat such as $(CGG)_n$ and $(CAG)_n$, which forms hairpin-like structure, potentially causes longer product than the template,(30) the transient hairpin structure of the nascent DNA product from cBcl2 template might increase the size of product. Therefore, we calculated the rate of full-length product formation including the smeared bands. As a result, the $\ln k_s$ value was -0.18, and $-\Delta G°_{37}$ value was 3.0 kcal mol^{-1} (SI Appendix, Figs. S15B,C). On the $\ln k_s$ vs. $-\Delta G°_{37}$ plot, the cBcl2 i-motif conformed to the linear correlation of Class I i-motifs (Fig. 3B), which indicates that the topology of i-motif does not influence the replication stall. Based on these experiments, we conclude that the Class II i-motifs with longer loops and Class I motifs are unwound by a similar mechanism.

The topology and stability of non-canonical structures are also affected by the crowding condition. Experiments were also performed in the presence of 20 wt% PEG200 (Figs. 3D, SI Appendix, Fig. S12). In the case of the i-motif, the slope of $\ln k_s$ vs $-\Delta G°_{37}$ was -0.12 (Fig. S16), whose magnitude was about six-fold smaller than that in the absence of PEG200 (-0.74) (SI Appendix, Fig. S14B). Thus, the stalling effect of i-motif decreased in 20 wt% PEG200 compared to conditions without PEG200. Our recent study suggested that ethylene glycol binds to the duplex and disrupts the hydrogen network around Watson-Crick base pairing.(31) All stabilities of the i-motif in the presence of 20 wt% PEG200 decreased compared to stabilities in corresponding conditions without PEG200. Therefore, PEG200 might uniquely interact with the base pairs of the i-motif and decrease stability, thereby increasing replication efficiency. On the other hand, in

the presence of 20 wt% PEG200 and 30 mM KCl induced the transformation of $(T_2AG_3)_4$ from a mixed to a parallel topology (SI Appendix, Fig. S3).(32) Replication was effectively repressed and showed a similar slope of the $\ln k_s$ vs $-\Delta G°_{37}$ plot (-0.87) to that of the i-motif observed in the absence of PEG200 (-0.83) (SI Appendix, Figs. S16, S14A). Thus, this transformation may inhibit replication more efficiently as observed in the case of ILPR. In the presence of PEG1000, the replication of i-motifs was effectively repressed (Fig. 3D). The slope of $\ln k_s$ vs $-\Delta G°_{37}$ was -0.56, which was close to that observed in the absence of PEG200 (-0.83) (SI Appendix, Figs. S16, S14C). Therefore, a specific molecular environment with various conditions of molecular crowding regulates the processivity of KF along a template DNA based on the activation free energy for unwinding by changing the stability and the topology of the DNA structure formed.

DISCUSSION

In this study, we found a linear correlation between the logarithm of the replication rate constant ($\ln k_s$) and the stability ($-\Delta G°_{37}$) of particular structural topologies of non-canonical structures in the template DNA, indicating that the rate limiting step of structure unwinding by KF is determined by both stability and topology of the structure formed within the DNA template. We reported previously that the processivity of T7 RNA polymerase in the transcription reaction was influenced by both hairpin and G4 structures formed within the template DNAs.(6) The rates of transcription reaction were correlated with the stabilities of the structures, and the lack of relationship with topology may indicate that the unwinding mechanisms of RNA polymerase and KF are different.(33) Some G4-specific helicases show topological specificity. For example, the telomere protein TPP1 more efficiently unwinds antiparallel G4 structures like Q6 than it does parallel G4.(34) In contrast, RNA helicase RHAU preferentially binds to parallel G4 structures like that of ILPR.(35) These results imply that the processivity of DNA polymerase requires the assistance of different types of helicases able to unwind various topologies.

KF has an ability to unwind hairpin structures, presumably through an unzipping mechanism, which is essentially the reverse of the folding process. Based on the nearest neighbor model, stability of a duplex region of a hairpin is determined by the sequence and number of stacked base pairs.(36, 37) Terminal base pairs “breathe” and are easily dissociated.(38) Therefore, KF may induce unzipping of a hairpin in a mechanism that does not depend on the overall free energy of the hairpin structure (Fig. 4A). In the case of i-motif, the structural topology

encountered by the polymerase is quite different from those in hairpin and mixed G4 structures, and the loop structure of the i-motif likely presents a steric obstruction to unwinding of $C:C^+$ base pairs (Fig. 4B). Furthermore, in the i-motif structure, two parallel strand $C:C^+$ base pairs mutually intercalate with each other and form the tetraplex. In this conformation, the base pair next to the terminal base pair belongs to the other parallel strand. Therefore, the consecutive unzipping from the terminal base pairs by breathing should be repressed due to the stacking by the intercalated base pairs from the other parallel strand. Replication through an i-motif-forming region may require the complete unfolding of the structure, leading to our observations. This might be a reason why G4 showed different enzyme processivities. In the case of mixed G4, although the stability of a terminal quartet is presumably higher than that of a terminal base pair of a duplex, once the terminal quartet is unzipped by breathing, the stability of the overall structure is decreased and KF can proceed (Fig. 4C). As for other G4 topologies, the breathing of the terminal quartet might be repressed, because metastable intermediates formed due to strand slippage via triplex formation occur during unfolding (Fig. 4D).(39) Data on the ILPR and Q6 G4s fall on the line of $\ln k_s$ vs $-\Delta G^\circ_{37}$ data on i-motif-forming sequences. This suggests that it is also possible that metastable intermediates in the unfolding reaction might prolong dissociation of the i-motif.

In cells, helicases assist DNA polymerase in replication of structured DNAs. A variety of helicases that unwind G4 have been identified, and defects in helicases that unwind these structures are associated with genetic diseases.(40) This phenomenon implies that the topological properties of structures formed in DNA templates are important. A recent study showed that the anti-parallel G4 on the leading strand with T_m values even less than physiological temperature could cause genomic instability.(41) This phenomenon implies that the replication stall due to the topology of non-canonical structures is also influenced by helicase polarity. Helicases that unwind i-motif structures have not yet been identified. However, i-motif binding proteins such as hnRNPA1 and hnRNPLL induce unwinding of i-motif structures.(42-44) It has also been suggested that negative supercoiling induced by RNA polymerases promote i-motif formation in cells.(45) Our data suggest that i-motif formation can block replication, and this in turn may result in genomic instability. Genomic instability is associated with cancer development and also with chronic diseases such as diabetes.(46, 47) It is possible that mutation in or dysregulation of expression of i-motif binding proteins or the acidic pH of the cancer cell environment may promote or lead to stabilization of i-motif formation. Moreover, there

exists a lower number of the free water in tumor cells than normal cells,(48) which might be due to highly crowded condition and extremely stabilize i-motif formation.

In conclusion, template DNA that forms an i-motif especially decreased the processivity of DNA polymerase through a mechanism likely due to the unique topology of base pairs that stabilize the i-motif. Unwinding of i-motif structure by processive activity of KF had an about 3-fold higher activation free energy barrier ($\Delta G°_{37}{}^{‡}$) than that for unwinding of mixed G4s. The stalling was correlated with stability of the i-motif structure and was regulated by crowding molecule size. Therefore, it is possible that the i-motif may cause genomic instability in cancer cells. Our novel physicochemical approach can be applied to the screening of the ligands capable of inducing formation of specific topologies that stall replication. Development of topology-specific binders may enable site-specific control of genomic instability and expression of genes triggered by the non-canonical structures targeted.

Materials and Methods

Detailed information of the materials and methods used in this study are provided in *SI Materials and Methods*.

Acknowledgements

This work was supported in part by Grants-in-Aid for Scientific Research and MEXT-Supported Program for the Strategic Research Foundation at Private Universities (2014-2019), Japan, The Hirao Taro Foundation of KONAN GAKUEN for Academic Research, The Okazaki Kazuo Foundation of KONAN GAKUEN for Advanced Scientific Research, The Chubei Itoh Foundation and Hyogo Science and Technology Association.

REFERENCES

1. Nakano S, Miyoshi D, & Sugimoto N (2014) Effects of molecular crowding on the structures, interactions, and functions of nucleic acids. *Chem. Rev.* 114(5):2733-2758.
2. Biffi G, Tannahill D, McCafferty J, & Balasubramanian S (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5(3):182-186.
3. Fadloun A*, et al.* (2013) Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat. Struct. Mol. Biol.* 20(3):332-338.
4. Bochman ML, Paeschke K, & Zakian VA (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.* 13(11):770-780.
5. Siddiqui-Jain A, Grand CL, Bearss DJ, & Hurley LH (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* 99(18):11593-11598.
6. Tateishi-Karimata H, Isono N, & Sugimoto N (2014) New insights into transcription fidelity: thermal stability of non-canonical structures in template DNA regulates transcriptional arrest, pause, and slippage. *PloS one* 9(3):e90580.
7. Endoh T, Kawasaki Y, & Sugimoto N (2013) Suppression of Gene Expression by G-Quadruplexes in Open Reading Frames Depends on G-Quadruplex Stability. *Angew. Chem. Int. Ed.* 125(21):5632-5636.
8. Bugaut A & Balasubramanian S (2012) 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.* 40(11):4727-4741.
9. Wu Y & Brosh RM (2010) G-quadruplex nucleic acids and human disease. *FEBS J.* 277(17):3470-3488.
10. Benabou S, Aviñó A, Eritja R, González C, & Gargallo R (2014) Fundamental aspects of the nucleic acid i-motif structures. *RSC Advances* 4(51):26956-26980.

11. Phan AT & Mergny JL (2002) Human telomeric DNA: G-quadruplex, i-motif and Watson–Crick double helix. *Nucleic Acids Res.* 30(21):4618-4625.
12. Brazier JA, Shah A, & Brown GD (2012) I-Motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.* 48(87):10739-10741.
13. Kendrick S & Hurley LH (2010) The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements. *Pure Appl. Chem.* 82(8):1609-1621.
14. Gehring K, Leroy J-L, & Guéron M (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* 363:561-565.
15. Phan AT, Guéron M, & Leroy J-L (2000) The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* 299(1):123-144.
16. Kendrick S*, et al.* (2014) The Dynamic Character of the BCL2 Promoter i-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J. Am. Chem. Soc.* 136(11):4161-4171.
17. Gurung SP, Schwarz C, Hall JP, Cardin CJ, & Brazier JA (2015) The importance of loop length on the stability of i-motif structures. *Chem. Commun.* 51(26):5630-5632.
18. Dhakal S*, et al.* (2012) G-quadruplex and i-motif are mutually exclusive in ILPR double-stranded DNA. *Biophys. J.* 102(11):2575-2584.
19. Brosh RM, Jr. (2013) DNA helicases involved in DNA repair and their roles in cancer. *Nature reviews. Cancer* 13(8):542-558.
20. Paeschke K*, et al.* (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* 497(7450):458-462.
21. Beese LS, Derbyshire V, & Steitz TA (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* 260(5106):352-355.
22. Chen H*, et al.* (2014) Exploring the formation and recognition of an important G-quadruplex in a HIF1alpha promoter and its transcriptional inhibition by a benzo[c]phenanthridine derivative. *J. Am. Chem. Soc.* 136(6):2583-2591.

23. Rajendran A, Nakano S-i, & Sugimoto N (2010) Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* 46(8):1299-1301.
24. Cui J, Waltman P, Le VH, & Lewis EA (2013) The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH. *Molecules* 18(10):12751-12767.
25. Howorka S, Movileanu L, Braha O, & Bayley H (2001) Kinetics of duplex formation for individual DNA strands within a single protein nanopore. *Proc. Natl. Acad. Sci. U. S. A.* 98(23):12996-13001.
26. Rauzan B*, et al.* (2013) Kinetics and thermodynamics of DNA, RNA, and hybrid duplex formation. *Biochemistry* 52(5):765-772.
27. Crothers DM, Bloomfield VA, & Tinoco I (2000) *Nucleic acids: structures, properties, and functions* (University science books).
28. Gu XB, Nakano S, & Sugimoto N (2007) Consecutive GC base pairs determine the energy barrier of DNA duplex formation under molecularly crowded conditions. *Chem. Commun.* (26):2750-2752.
29. Yu Z*, et al.* (2009) ILPR G-quadruplexes formed in seconds demonstrate high mechanical stabilities. *J. Am. Chem. Soc.* 131(5):1876-1882.
30. Mirkin EV & Mirkin SM (2014) To switch or not to switch: at the origin of repeat expansion disease. *Mol. Cell* 53(1):1-3.
31. Nakano M*, et al.* (2015) Thermodynamic properties of water molecules in the presence of cosolute depend on DNA structure: a study using grid inhomogeneous solvation theory. *Nucleic Acids Res.* 43(21):10114-10125.
32. Yu H, Gu X, Nakano S, Miyoshi D, & Sugimoto N (2012) Beads-on-a-string structure of long telomeric DNAs under molecular crowding conditions. *J. Am. Chem. Soc.* 134(49):20060-20069.
33. Yin YW & Steitz TA (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science* 298(5597):1387-1395.
34. Ray S, Bandaria JN, Qureshi MH, Yildiz A, & Balci H (2014) G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. *Proc. Natl. Acad. Sci. U. S. A.* 111(8):2990-2995.
35. Heddi B, Cheong VV, Martadinata H, & Phan AT (2015) Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution

structure of a peptide–quadruplex complex. *Proc. Natl. Acad. Sci. U. S. A.* 112(31):9608-9613.

36. Turner DH, Sugimoto N, Kierzek R, & Dreiker SD (1987) Free energy increments for hydrogen bonds in nucleic acid base pairs. *J. Am. Chem. Soc.* 109(12):3783-3785.
37. Turner DH, Sugimoto N, & Freier SM (1988) RNA structure prediction. *Annu. Rev. Biophys. Biophys. Chem.* 17:167-192.
38. Phelps C, Lee W, Jose D, von Hippel PH, & Marcus AH (2013) Single-molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions. *Proc. Natl. Acad. Sci. U. S. A.* 110(43):17320-17325.
39. Stadlbauer P, Krepl M, Cheatham TE, 3rd, Koca J, & Sponer J (2013) Structural dynamics of possible late-stage intermediates in folding of quadruplex DNA studied by molecular simulations. *Nucleic Acids Res.* 41(14):7128-7143.
40. Wu Y & Brosh RM (2010) G-quadruplex nucleic acids and human disease. *FEBS J.* 277(17):3470-3488.
41. Schiavone D*, et al.* (2014) Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. *EMBO J.* 33(21):2507-2520.
42. Kang HJ, Kendrick S, Hecht SM, & Hurley LH (2014) The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* 136(11):4172-4185.
43. Roy B*, et al.* (2016) Interaction of Individual Structural Domains of hnRNP LL with the BCL2 Promoter i-Motif DNA. *J. Am. Chem. Soc.* 138(34):10950-10962.
44. Miglietta G, Cogoi S, Pedersen EB, & Xodo LE (2015) GC-elements controlling HRAS transcription form i-motif structures unfolded by heterogeneous ribonucleoprotein particle A1. *Scientific reports* 5:18097.
45. Brooks TA & Hurley LH (2009) The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat. Rev. Cancer* 9(12):849-861.
46. Palazzo RP, Bagatini PB, Schefer PB, de Andrade FM, & Maluf SW (2012) Genomic instability in patients with type 2 diabetes mellitus on hemodialysis. *Rev. Bras. Hematol. Hemoter.* 34(1):31-35.

47. Gaillard H, Garcia-Muse T, & Aguilera A (2015) Replication stress and cancer. *Nat. Rev. Cancer* 15(5):276-289.
48. Damadian R, Zaner K, Hor D, & DiMaio T (1974) Human tumors detected by nuclear magnetic resonance. *Proc. Natl. Acad. Sci. U. S. A.* 71(4):1471-1473.

Table 1. Thermodynamic and kinetic parameters of the structured sequence region on each template DNA at pH 6.0

DNA topology	Sequence	T_m (°C)	$-\Delta G°_{37}$ (kcal mol^{-1})	k_s (min^{-1})
i-motif	$(C_3TA_2)_4$	31.4± 0.1	-0.70 ± 0.01	4.5 ± 0.1
	Hif1a	57.3 ± 0.1	3.1 ± 0.2	0.39 ± 0.11
Hairpin	H2	75.8 ± 2.2	4.0 ± 0.2	3.7 ± 0.1
Mixed G4[b]	$(T_2AG_3)_4$	62.8 ± 0.5	3.0 ± 0.2	2.6 ± 0.3
Parallel G4	ILPR	58.1 ± 0.2	2.1 ± 0.3	0.54 ± 0.4
Anti-parallel G4	Q6	71.1 ± 0.3	4.5 ± 0.1	0.08 ± 0.03

[a] All experiments were performed in 40 mM MES (pH 6.0) and 8 mM $MgCl_2$ without both KCl and PEG except for the case of $(T_2AG_3)_4$. Oligonucleotides corresponding to the structured region of indicated templates were evaluated at 10 μM strand concentration.

[b] Experiments for $(T_2AG_3)_4$ were performed in 40 mM MES (pH 6.0), 8 mM $MgCl_2$, and 30 mM KCl.

Figures

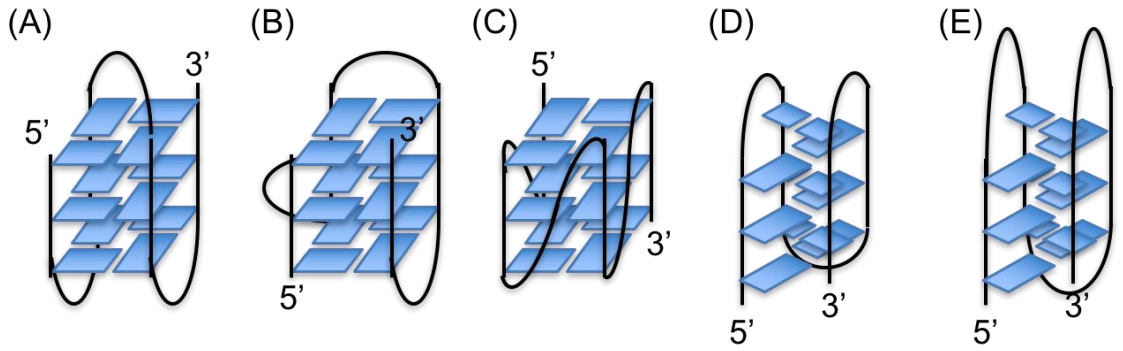


Figure 1. Schematic illustrations of topologies of (A) antiparallel G4, (B) mixed G4, (C) parallel G4, (D) Class I i-motif and (E) Class II i-motif having longer loops than Class I.

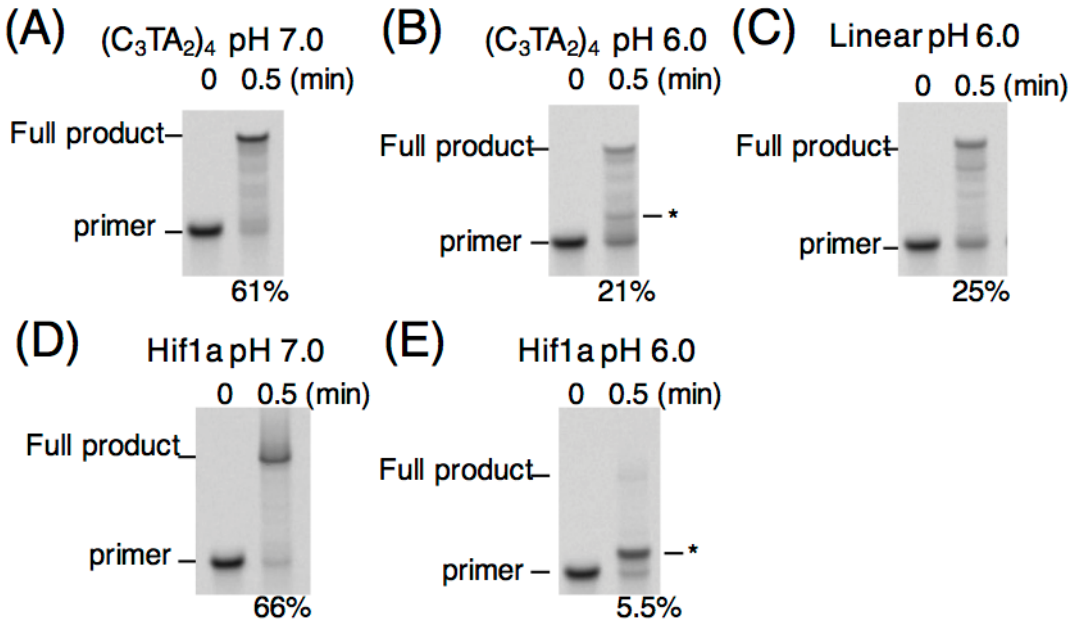


Figure 2. Representative PAGE analyses of replication reactions with (A) $(C_3TA_2)_4$ template DNA at pH 7.0, (B) $(C_3TA_2)_4$ template DNA at pH 6.0, (C) linear template DNA at pH 6.0, (D) Hif1a template DNA at pH 7.0 and (E) Hif1a template DNA at pH 6.0. The images were captured using a fluorescent imager, thus only DNA containing the fluorescent label is visualized. Images of the same gels stained to visualize all DNA are shown in Fig. S2. Reactions were carried out with 1

μM KF, 1 μM DNAs and 250 μM dNTPs in 40 mM MES buffer (pH 7.0 or 6.0) for 2 min at 37 °C. The asterisk indicates the position of major stalled product. The values at the bottom of the gel image are the percentages of full-length product.

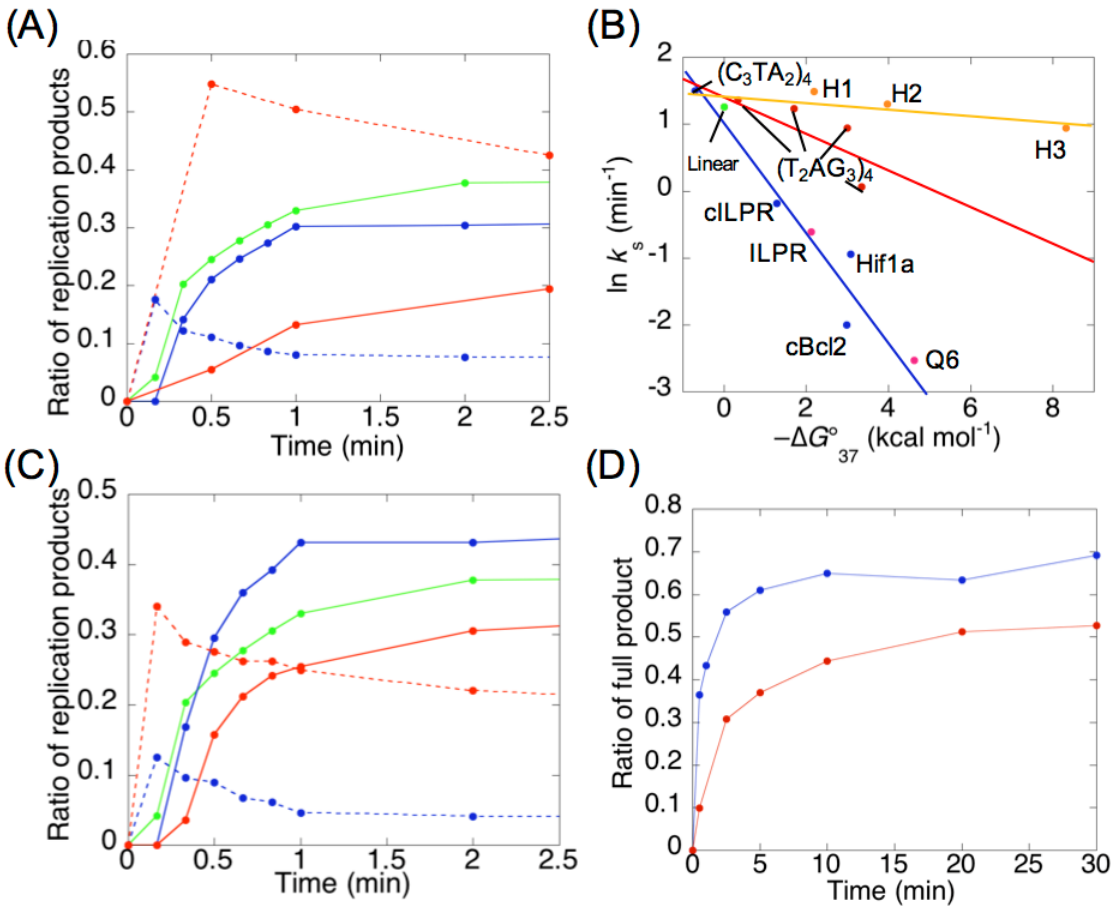


Figure 3. Kinetics of replication of DNA templates containing non-canonical structures. (A) Ratios of stalled (dotted lines) and full-length (solid lines) to total product of $(C_3TA_2)_4$ (blue), Hif1a (red), and linear (green) templates as a function of time. (B) Plot of $-\Delta G°_{37}$ values versus the logarithms of rate constants (k_s) for reactions to dissolve the stall from reaction start along i-motif-forming templates (blue), G4-forming templates (red), and linear template (green). Data on G4 ILPR template was excluded from fitting of the G4 data. (C) Ratios of stalled and full-length to total products of $(T_2AG_3)_4$ replication in the presence of 1 mM KCl (blue) and 30 mM KCl (red) and of linear (green) replication in the absence of KCl as a function of time. (D) Ratios of full-length product to total product of Hif1a in the presence of 100 mM KCl and 20 wt% PEG200 (blue) and in the presence of 100 mM KCl and 20 wt% PEG1000 (red). All the reactions

were carried in 40 mM MES (pH 6.0), 8 mM $MgCl_2$, 1 μM KF, 1 μM DNAs, and 250 μM dNTPs at 37 °C.

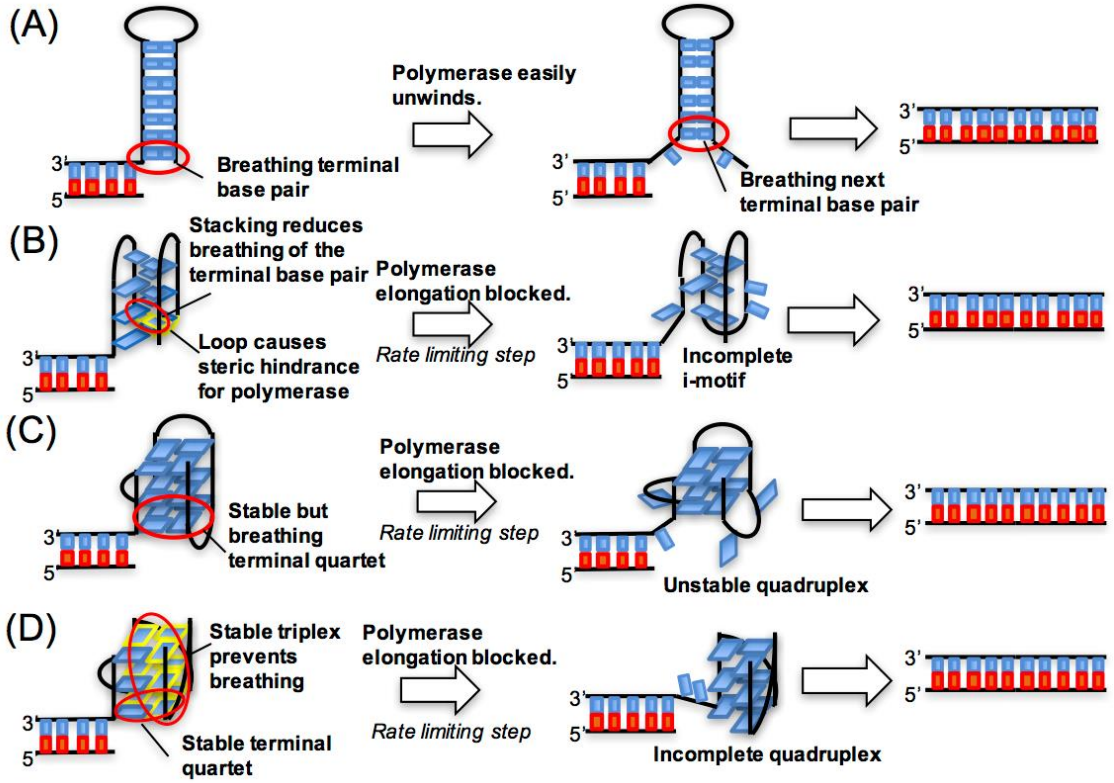


Figure 4. Proposed mechanism of unwinding of structured DNAs in templates containing various structures. (A) The terminal base pair of a hairpin is relatively unstable. Thus, polymerase can induce unzipping one base pair at a time to proceed through a hairpin. (B) In the i-motif structure, the loop region can be a block for polymerase, and the base pair has a different topology from the first, which likely stalls the polymerase. (C) Replication of mixed G4 structure presents an obstacle as the terminal quartet is more stable than the terminal base pair of a hairpin. Once the terminal G-quartet is unzipped, the stability of G4 is significantly reduced and replication proceeds rapidly. (D) Parallel/anti-parallel G4s have a different unfolding pathway, which repressed the terminal breathing and enzyme processivity as well as the i-motif.

Supporting Information

Topological impact of non-canonical DNA structures on Klenow Fragment of DNA polymerase

Shuntaro Takahashi[1], John A. Brazier[2], and Naoki Sugimoto[1,3*]

[1] Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Kobe 650-0047, Japan

[2] Department of Pharmacy, University of Reading, Whiteknights, Reading, Berks RG6 6AD, UK

[3] Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Kobe 650-0047, Japan

* To whom correspondence should be addressed. Tel: (+81)78-303-1457; Fax: (+81)78-303-1495; Email: sugimoto@konan-u.ac.jp

Materials and Methods

Materials. dNTPs were purchased from Takara Bio. Other reagents were purchased from Wako Pure Chemicals and used without further purification.

Oligonucleotides. HPLC-purified FAM-labeled primer and template DNAs were purchased from Japan Bio Service. All DNA sequences used in this study for replication assays are listed in Table 1.

Enzyme. The gene encoding Klenow Fragment was amplified from *E.coli* JM109 genome DNA by PCR carried out with PrimeSTAR DNA polymerase (Takara Bio) and primers (5'-GGGACCATATGGTGATTTCTTATGACAACTACG-3' and 5'-GGGAGAATTCTTAGTGCGCCTGATCCCAG-3') purchased from Eurofin Genomics. The cloned DNA fragments were digested with NdeI and EcoRI, and cloned into pMal-p5x vector (New England Bio Labs). *E.coli* EG2523 (New England Bio Labs) was transformed with the constructed vector. The cells were cultured in LB medium to an A_{600} of around 0.5, followed by addition of isopropyl β-D-1-thiogalactopyranoside and further culture. The cultured cells were harvested and lysed. The soluble fraction was loaded on the column packed with amylose resin (New England BioLabs). After treatment with the Factor Xa protease, KF was purified over a Hitrap Heparin column followed by purification through Hiload Superdex 200 (GE Healthcare). The concentration was determined by UV absorbance at 280 nm using the molar extinction coefficient of 58,790.

DNA polymerase gp5 from T7 phage conjugated with *E. coli* thioredoxin was purchased from New England BioLabs. The helicase gp4 was overexpressed from the *E. coli*

BL21(DE3) pLysS strain (Merck) carrying the expression plasmid encoding the C-terminus his-tagged gp4 cloned from T7 DNA. The recombinant gp4 was purified through Histrap HP, Hitrap Heparin, and Hiload Superdex 200 columns.

Replication assay. Primer and template DNA were annealed in the buffer used in the replication reaction: 40 mM MES (pH 6.0), 8 mM $MgCl_2$, 1 μM KF, 1 μM DNA, and 250 μM dNTPs. KCl and PEG200 or PEG1000 were added as indicated. After preparation of the solution, the mixtures were incubated at 37 °C. The reaction was stopped by addition of the solution containing EDTA and formamide. Products were separated by on 12% polyacrylamide gels containing 8 M urea at 70 °C for 1 hour at 200 V in TBE buffer. The gel images were captured using a Fujifilm FLA-5100 fluorescent imager before and after staining with SYBR Gold (ThermoFisher Scientific). The intensities of bands were analyzed by NIH ImageJ software. The amount of full-length product (*P*) was quantified by calculating the ratio of intensity of full-length product band to intensity of all bands. The kinetic model was applied to the two-step sequential model;

$$P_0 \overset{k_s}{\rightarrow} P_S \overset{k_f}{\rightarrow} P_F$$

where P_0 is the starting state of reaction, P_s is the state immediately after dissolving the stall by non-canonical structure is removed, P_f is the state after finishing replication fully, k_s (min^{-1}) is the rate constant of dissolving the stall from reaction start by KF, and k_f (min^{-1}) is the rate constant of full-length product after dissolving replication stall. Rate constants were evaluated by global fit using Dynafit (Biokin) and linear correlations were analyzed by KaleidaGraph (Synergy Software). All experiments were carried out in three independent experiments at least.

Circular dichroism measurements. For CD measurements, 10 μM oligonucleotide was dissolved in buffer used in the replication assay without dNTPs. Samples were

heated at 95 °C for 3 min and then cooled to 20 °C at the rate of 1.0 °C/min. The CD spectra were collected using a JASCO J-1500 at 37 °C. We assigned the structures from CD spectra from the typical characteristics (i-motif: a positive peak at 290 nm, anti-parallel G4: a positive peak at 295 nm, parallel G4: a positive peak at 265 nm, mixed: a spectra like mixture of anti-parallel and parallel G4).

UV melting assay. For melting analyses, 10 μM oligonucleotide with the sequence of a structure forming region within the template (Table 1) was dissolved in buffer used in the replication assay without dNTPs. The melting analyses were carried out on the Shimadzu UV-1800 equipped with a temperature control system. Samples were cooled from 90 °C to 0 °C at the rate of 1.0 °C/min and then temperature was increased from 0 °C to 90 °C at the rate of 0.5 °C/min. The UV melting curves were normalized and analyzed by curve fitting by KaleidaGraph (Synergy Software) to determine thermodynamic parameters based on two-state model thermodynamics as described previously.(1) For i-motif sequences, we analyzed the change in absorbance at 295 nm based on a two-state model. Because the absorbance change at 295 nm showed the formation and dissociation of $C:C^+$ base parings, we considered the thermodynamic core stability of $C:C^+$ base pairs as the thermodynamic stability of the whole structure of i-motif.(2) All experiments were carried out in three independent experiments at least.

Reference

1. Takahashi S & Sugimoto N (2013) Effect of pressure on the stability of G-quadruplex DNA: thermodynamics under crowding conditions. *Angew. Chem. Int. Ed.* 52(51):13774-13778.
2. Mergny J L & Lacroix L (1998) Kinetics and thermodynamics of i-DNA formation: phosphodiester versus modified oligodeoxynucleotides. 26(21):4797-803.

Table S1. DNA sequences used in this study

	DNA sequence
FAM labeled primer	5'-FAM-GGGATGATACTAGTGCTTCGGCTTAATACGACTCACTATAGGG-3'
$(C_3TA_2)_4$ template	5'-CCCTAACCCTAACCCTAACCCTAACGGCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3'
$(T_2AG_3)_4$ template	5'-TTAGGGTTAGGGTTAGGGTTAGGGCGGCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3'
Hif1a template	5'-CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGCCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3’
Q6 template	5'-GGGGTTGGGGTGTGGGGTTGGGGAGGACACGGTGACCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3’
cILPR template	5’-CCCCACACCCCTGTCCCCACACCCCCGGCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3'
ILPR template	5’-GGGGTGTGGGGACAGGGGTGTGGGGCGGCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3’
H3 template	5'-GCCGTCCAACTATCGGACTTCGGTCCGATAGTTGGCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3’
H2 template	5'-GCCGTTTCGTAGTCTATCGGACTTCGGTCCGATAGCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3’
H1 template	5'-GCCGTTTCGTAGTATTTCTATCCGGACTTCGGTCCCCCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3’
Linear template	5'-GATTGATGTGATTGATTTGATTGATGTGATTGACCCTATAGTGAGTCGTATTAAGCCGAAGCACTAGTATCATCCC-3'
$(C_3TA_2)_4$	5’-CCCTAACCCTAACCCTAACCCTAA-3’

$(T_2AG_3)_4$	5'-TTAGGGTTAGGGTTAGGGTTAGGG-3’
Hif1a	5’-CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGC-3’
Q6	5’-GGGGTTGGGGTGTGGGGTTGGGG-3’
cILPR	5'-CCCCACACCCCTGTCCCCACACCCC-3’
ILPR	5'-GGGGTGTGGGGACAGGGGTGTGGGG-3’
H3	5'-CCAACTATCGGACTTCGGTCCGATAGTTGG-3’
H2	5'-CTATCGGACTTCGGTCCGATAG-3’
H1	5'-GGACTTCGGTCC-3’

Table S2. Thermodynamic and kinetic parameters of the structured sequence region on each template DNA

DNA Sequence	Condition	T_m (°C)	$-\Delta G°_{37}$ (kcal mol^{-1})	k_s (min^{-1})	k_f (min^{-1})
$(C_3TA_2)_4$	0 mM KCl and 0% PEG at pH 6.0	31.4 ± 0.1	-0.70 ± 0.01	4.5 ± 0.1	4.5 ± 0.1
	100 mM KCl and 0% PEG at pH 6.0	30.6 ± 0.1	-0.90 ± 0.03	3.6 ± 0.1	3.6 ± 0.1
	100 mM KCl and 20% PEG200 at pH 6.0	28.7 ± 0.1	-1.4 ± 0.1	2.8 ± 0.2	4.1 ± 0.9
	100 mM KCl and 20% PEG1000 at pH 6.0	37.3 ± 0.2	0.04 ± 0.03	1.2 ± 0.3	5.4 ± 1.8
cILPR	0 mM KCl and 0% PEG at pH 6.0	44.7 ± 0.1	1.3 ± 0.1	0.83 ± 0.14	11 ± 2.7
	100 mM KCl and 0% PEG at pH 6.0	45.4 ± 0.3	1.3 ± 0.1	0.78 ± 0.02	6.7 ± 2.0
	100 mM KCl and 20% PEG200 at pH 6.0	44.2 ± 0.1	1.0 ± 0.1	2.1 ± 0.4	9.0 ± 5.2
	100 mM KCl and 20% PEG1000 at pH 6.0	50.6 ± 0.2	2.0 ± 0.1	0.44 ± 0.02	7.0 ± 0.2
Hif1a	0 mM KCl and 0% PEG at pH 6.0	57.3 ± 0.1	3.1 ± 0.2	0.39 ± 0.11	6.3 ± 2.0
	100 mM KCl and 0% PEG at pH 6.0	54.9 ± 0.2	3.1 ± 0.1	0.072 ± 0.010	1.0 ± 0.1
	100 mM KCl and 20% PEG200 at pH 6.0	51.2 ± 0.1	3.1 ± 0.1	0.47 ± 0.03	15.2 ± 9.8
	100 mM KCl and 20% PEG1000 at pH 6.0	59.7 ± 0.2	4.2 ± 0.1	0.048 ± 0.009	0.86 ± 0.13
	100 mM KCl and 20% PEG1000 at pH 7.0	43.5 ± 0.1	0.8 ± 0.1	2.3 ± 0.8	4.0 ± 0.9
cBcl2	0 mM KCl and 0% PEG at pH 6.0	51.8 ± 0.1	3.0 ± 0.1	0.13 ± 0.02	1.9 ± 0.6
	100 mM KCl and 0% PEG at pH 6.0	51.2 ± 0.3	2.4 ± 0.2	0.55 ± 0.03	6.0 ± 2.0
	100 mM KCl and 20% PEG200 at pH 6.0	50.8 ± 0.1	2.3 ± 0.3	2.1 ± 0.4	9.0 ± 5.3

	100 mM KCl and 20% PEG1000 at pH 6.0	56.1 ± 0.1	3.8 ± 0.2	0.51 ± 0.12	1.8 ± 0.6
H1	0 mM KCl and 0% PEG at pH 6.0	55.8 ± 0.2	2.2 ± 0.1	4.4 ± 0.3	4.4 ± 0.3
H2	0 mM KCl and 0% PEG at pH 6.0	75.8 ± 2.2	4.0 ± 0.2	3.7 ± 0.1	3.7 ± 0.1
H3	0 mM KCl and 0% PEG at pH 6.0	77.4 ± 0.1	8.3 ± 0.8	2.6 ± 0.4	2.6 ± 0.4
$(T_2AG_3)_4$	1 mM KCl and 0% PEG at pH 6.0	41.5 ± 0.9	0.34 ± 0.07	3.9 ± 0.2	3.9 ± 0.2
	10 mM KCl and 0% PEG at pH 6.0	55.5 ± 0.2	1.7 ± 0.1	3.5 ± 0.1	3.5 ± 0.1
	30 mM KCl and 0% PEG at pH 6.0	62.8 ± 0.5	3.0 ± 0.2	2.6 ± 0.3	3.5 ± 0.4
	50 mM KCl and 0% PEG at pH 6.0	64.7 ± 0.4	3.3 ± 0.1	1.1 ± 0.4	4.0 ± 1.1
	1 mM KCl and 20% PEG200 at pH 6.0	58.1 ± 0.2	2.2 ± 0.1	0.48 ± 0.13	7.1 ± 3.2
	30 mM KCl and 20% PEG200 at pH 6.0	74.1 ± 1.0	4.0 ± 0.3	0.064 ± 0.002	1.0 ± 0.1
ILPR	ILPR	58.1 ± 0.2	2.1 ± 0.3	0.54 ± 0.4	2.7 ± 1.7
Q6	Q6	71.1 ± 0.3	4.5 ± 0.1	0.08 ± 0.03	1.2 ± 0.1
Linear	0 mM KCl and 0% PEG at pH 6.0	n.d.	n.d.	3.5 ± 0.4	6.7 ± 2.2
	100 mM KCl and 0% PEG at pH 6.0	n.d.	n.d.	1.4 ± 0.1	1.4 ± 0.1
	100 mM KCl and 20% PEG200 at pH 6.0	n.d.	n.d.	3.3 ± 0.1	3.3 ± 0.1
	100 mM KCl and 20% PEG1000 at pH 6.0	n.d.	n.d.	1.5 ± 0.2	2.1 ± 0.4
	0 mM KCl and 0% PEG at pH 7.0	n.d.	n.d.	18 ± 2.4	18 ± 2.4
	100 mM KCl and 20% PEG1000 at pH 6.0	n.d.	n.d.	3.6 ± 0.4	5.1 ± 0.8

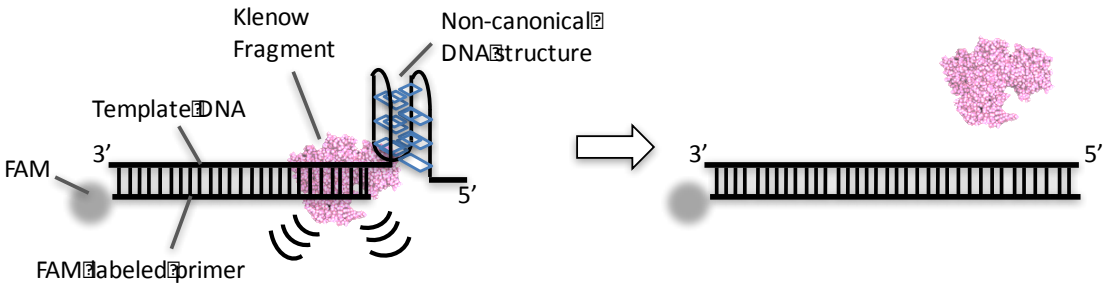
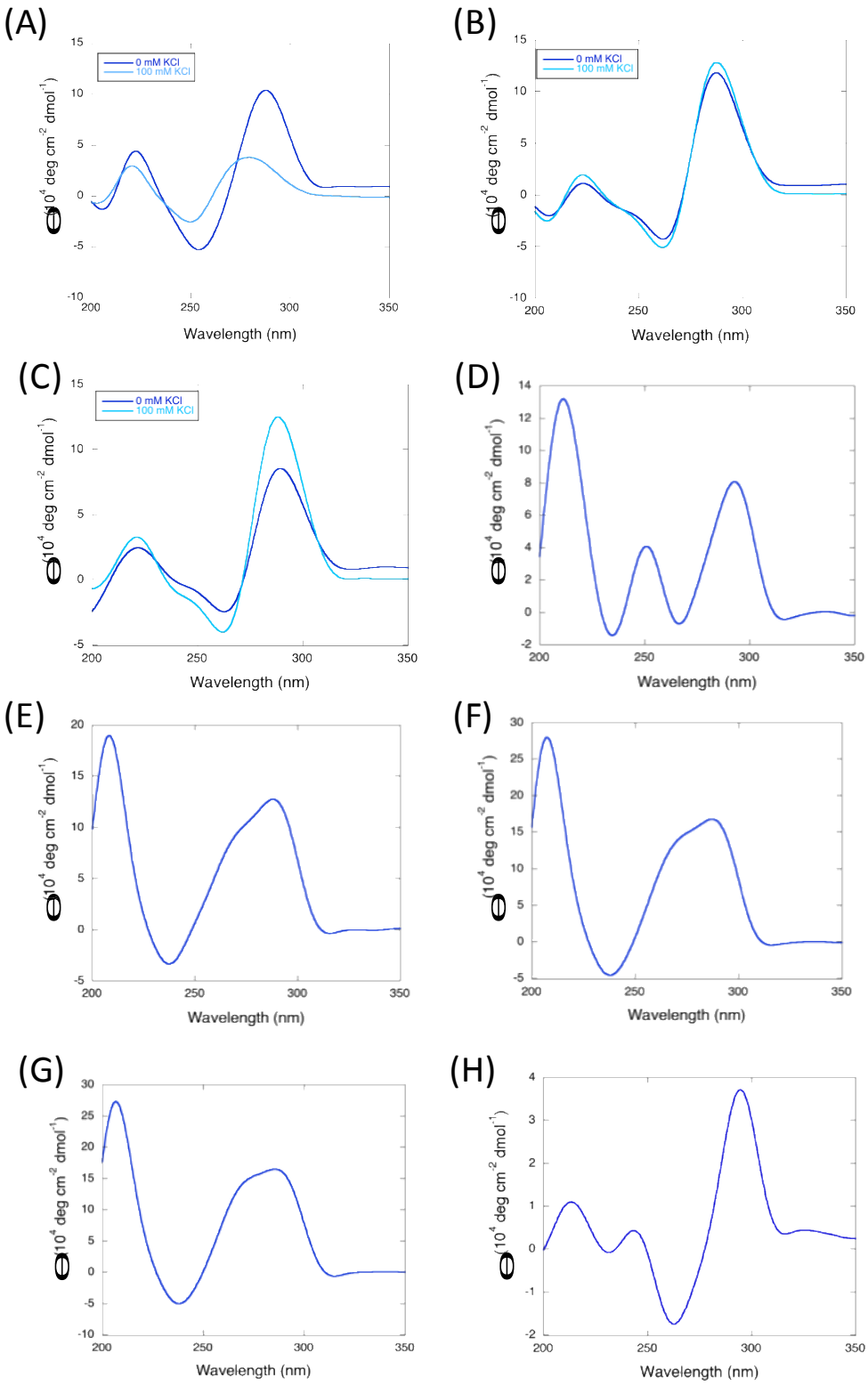


Figure S1. Schematic illustration of DNA replication assay. A FAM-labeled primer hybridized to a template is extended by KF. A region with the potential to form a non-canonical structure is located at four bases from the primer binding site. Product formation was quantified by measuring the fluorescence intensity of bands on denaturing PAGE.


(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
0 mM KCl
100 mM KCl
θ (10^4 deg cm^-2 dmol^-1)
Wavelength (nm)

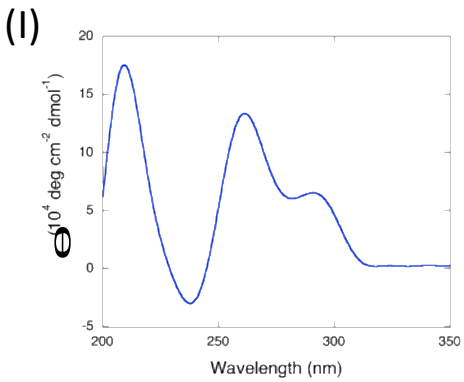


Figure S2. CD spectra of 10 µM oligonucleotides in the absence of PEG (A) $(C_3TA_2)_4$ in the absence of KCl (blue) and presence of 100 mM KCl (light blue), (B) Hif1a in the absence of KCl (blue) and presence of 100 mM KCl (light blue), (C) cILPR in the absence of KCl (blue) and presence of 100 mM KCl (light blue), (D) $(T_2AG_3)_4$ in the presence of 1 mM KCl, (E) $(T_2AG_3)_4$ in the presence of 10 mM KCl, (F) $(T_2AG_3)_4$ in the presence of 30 mM KCl, (G) $(T_2AG_3)_4$ in the presence of 50 mM KCl, (H) Q6 in the presence of 1 mM KCl, and (I) ILPR in the absence of KCl. Oligonucleotide sequences are shown in Table 1. All solutions were buffered with 40 mM MES (pH 6.0) and contained 8 mM $MgCl_2$ in the absence of PEG.

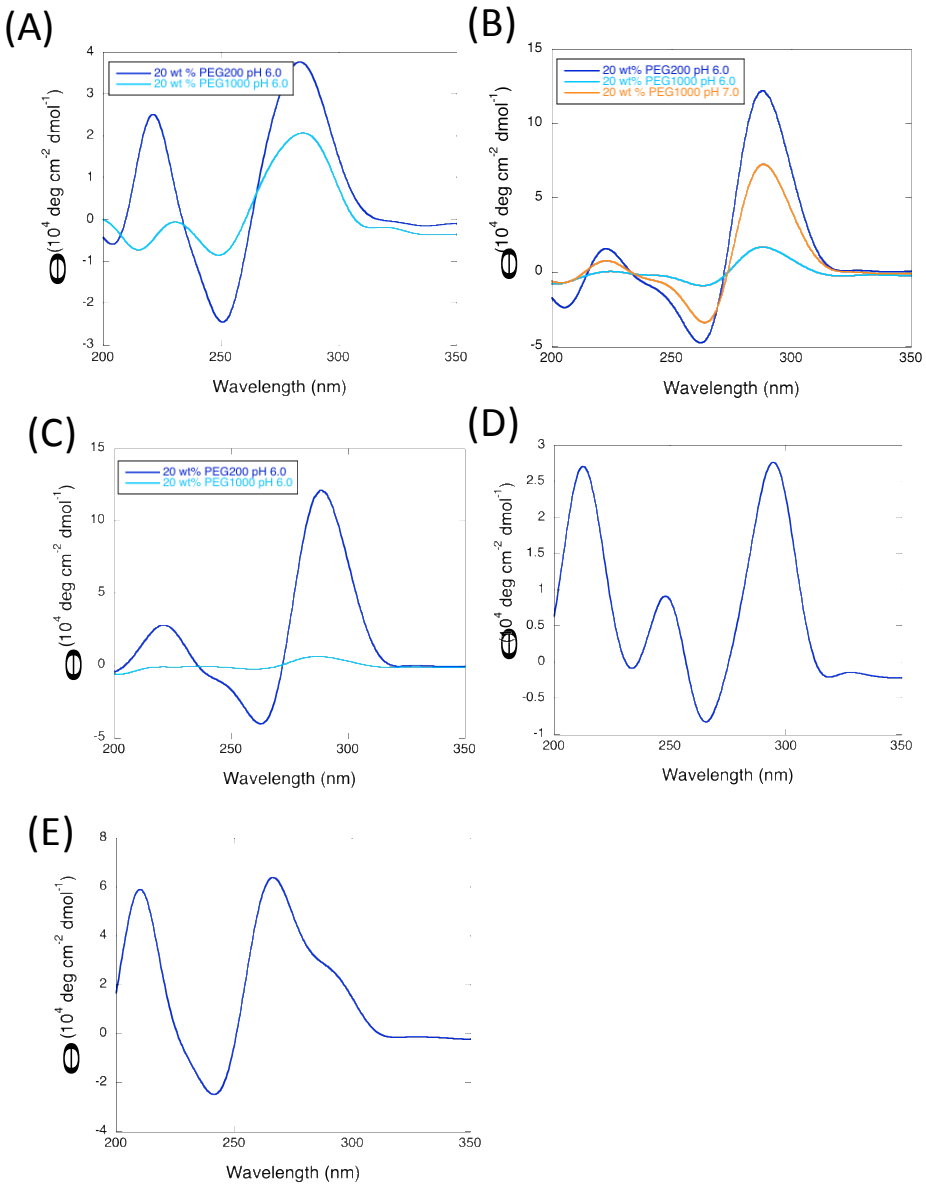


Figure S3. CD spectra of 10 µM oligonucleotides in the presence of PEG (A) $(C_3TA_2)_4$ (B) Hif1a, (C) cILPR, (D) $(T_2AG_3)_4$ in the presence of both 1 mM KCl and 20 wt% PEG200, and (E) $(T_2AG_3)_4$ in the presence of both 30 mM KCl and 20 wt% PEG200. The data obtained in the presence of 20 wt% PEG200 at pH 6.0, 20 wt% PEG1000 at pH 6.0, and 20 wt% PEG1000 at pH 7.0 are indicated in blue, light blue, and orange, respectively. Oligonucleotide sequences are shown in Table 1. All solutions were buffered with 40 mM MES (pH 6.0 or 7.0) and contained 100 mM KCl and 8 mM $MgCl_2$.

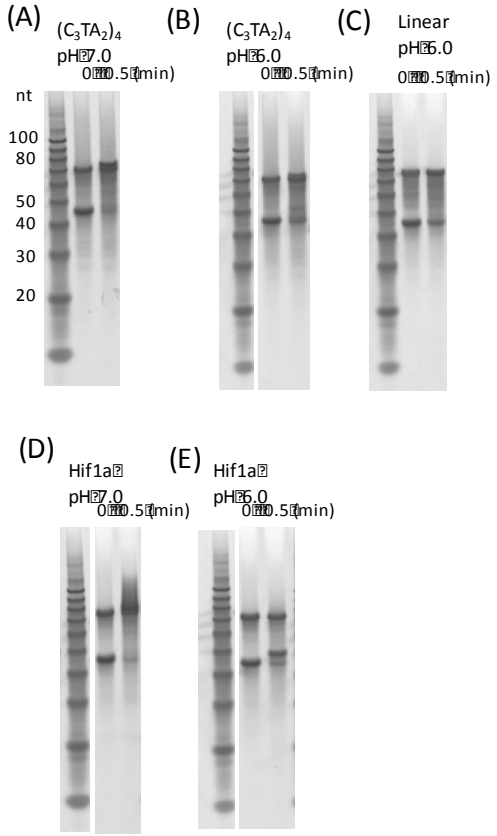


Figure S4. Images of gels shown in Fig. 2 stained with SYBR Gold reagent. PAGE analyses of replication reactions with (A) $(C_3TA_2)_4$ template DNA at pH 7.0, (B) $(C_3TA_2)_4$ template DNA at pH 6.0, (C) linear template DNA at pH 6.0, (D) Hif1a template DNA at pH 7.0 and (E) Hif1a template DNA at pH 6.0. In these images fluorescently labeled primer and products and unlabeled template strand are visible. Left-most lanes are 10-bp DNA ladders. In lanes of 0 min samples, upper bands correspond to the template DNA and lower bands to the FAM-labeled primer. At 0.5 min, the full-length product overlaps with the template DNA, and stalled product is observed just above the primer.

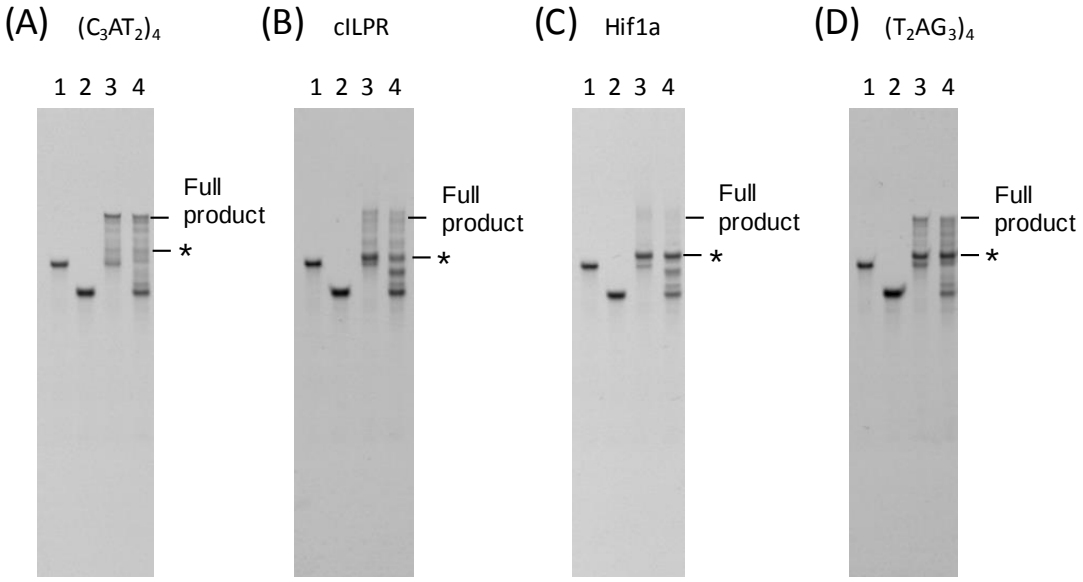


Figure S5. Effect of length of the primer on replication reactions with (A) $(C_3TA_2)_4$ template DNA, (B) cILPR template DNA, (C) Hif1a template DNA, and (D) $(T_2AG_3)_4$ template DNA . Lane 1: FAM labeled primer; Lane 2: FAM-labeled primer S; Lane 3: after 0.5 min reaction in the presence of FAM labeled primer; Lane 4: after 0.5 min reaction in the presence of FAM labeled primer S. The asterisk on the right of each gel image indicates the migrated position of major stalled products. All the reactions were carried in 40 mM MES (pH 6.0), 8 mM $MgCl_2$, 30 mM KCl, 1 μM KF, 1 μM DNA template and primer, and 250 μM dNTPs at 37 °C.

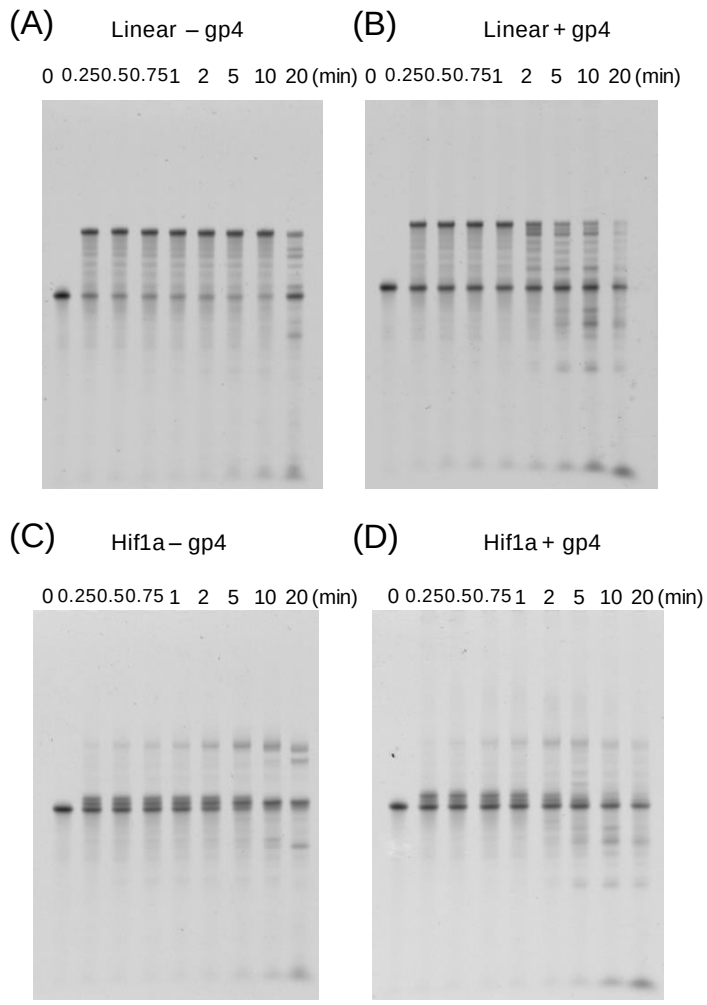
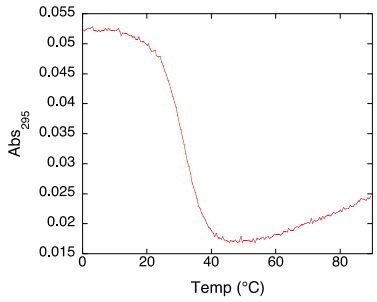


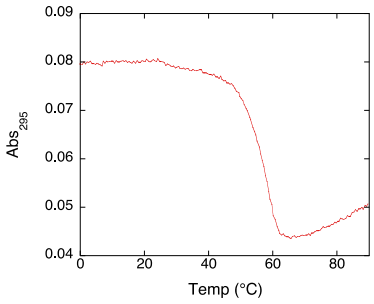
Figure S6. PAGE images of the replication by T7 DNA polymerase of (A) Linear template DNA without gp4, (B) Linear template DNA with gp4, (C) Hif1a template DNA without gp4 and (D) Hif1a template DNA with gp4. Degradation is due to the exonuclease activity of T7 DNAP system. Since T7 DNA polymerase requires gp4 DNA helicase for the processive replication of a duplex, the replication was performed in the presence and absence of gp4. In the case of Hif1a, the stalling of replication was comparable to that by KF, and amount of the stalled product was not influenced by the presence of gp4. These results indicate that DNA polymerases from different organisms were stalled by the i-motif. All the reactions were carried in 40 mM MES (pH 6.0), 8 mM

$MgCl_2$, 1 µM T7 DNA polymerase conjugated to *E. coli* thioredoxin, 0 or 6 µM gp4, 1 µM DNA template and primer, and 250 µM dNTPs at 37 °C.

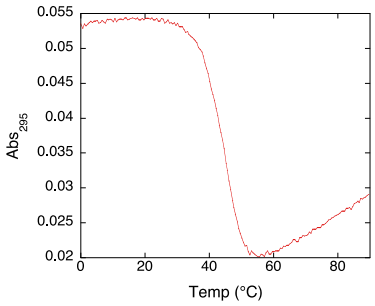
(A)


Abs295
Temp (°C)

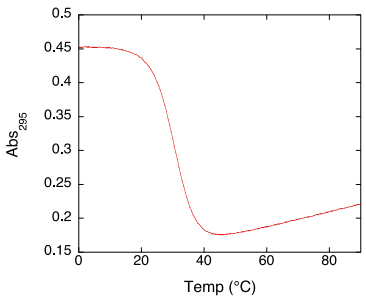
(B)


Abs295
Temp (°C)

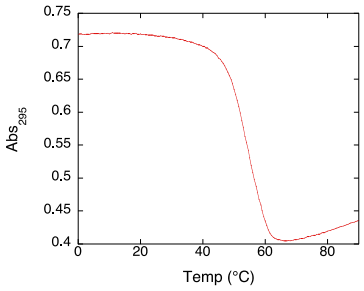
(C)


Abs295
Temp (°C)

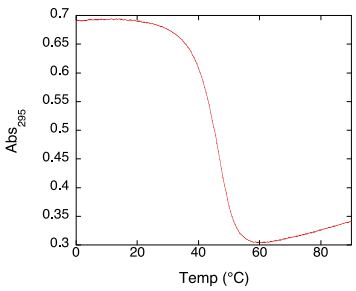
(D)


Abs295
Temp (°C)

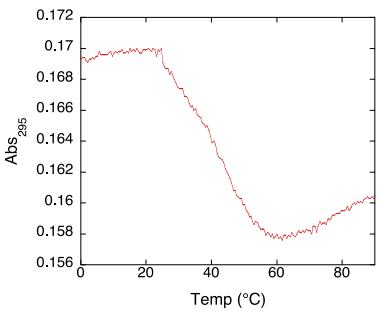
(E)


Abs295
Temp (°C)

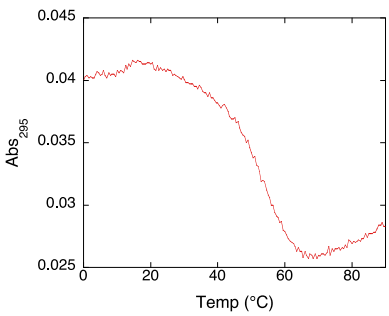
(F)


Abs295
Temp (°C)

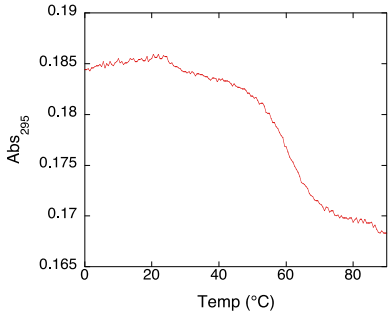
(G)


Abs295
Temp (°C)

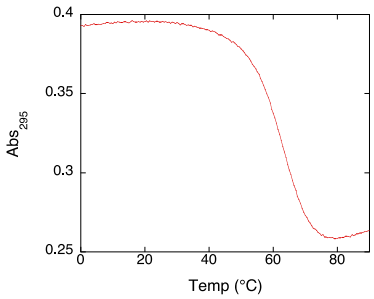
(H)


Abs295
Temp (°C)

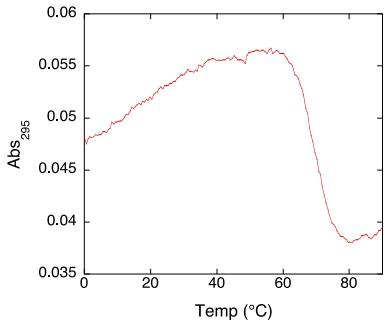
(I)


0.19
0.185
0.18
0.175
0.17
0.165
Abs295
0
20
40
60
80
Temp (°C)

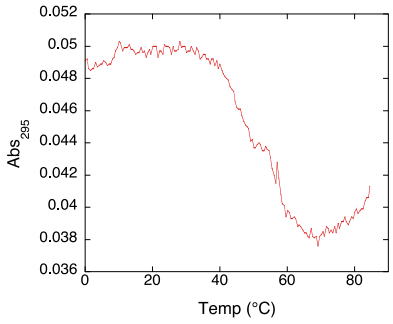
(J)


0.4
0.35
0.3
0.25
Abs295
0
20
40
60
80
Temp (°C)

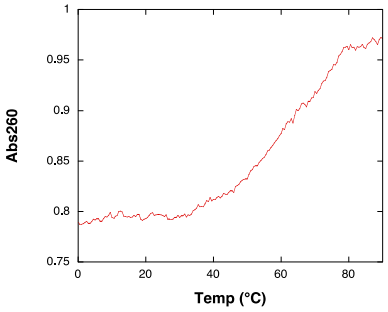
(K)


0.06
0.055
0.05
0.045
0.04
0.035
Abs295
0
20
40
60
80
Temp (°C)

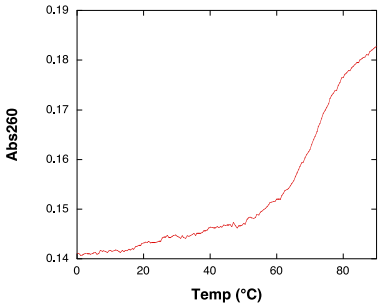
(L)


0.052
0.05
0.048
0.046
0.044
0.042
0.04
0.038
0.036
Abs295
0
20
40
60
80
Temp (°C)

(M)


1
0.95
0.9
0.85
0.8
0.75
Abs260
0
20
40
60
80
Temp (°C)

(N)


0.19
0.18
0.17
0.16
0.15
0.14
Abs260
0
20
40
60
80
Temp (°C)

(O)

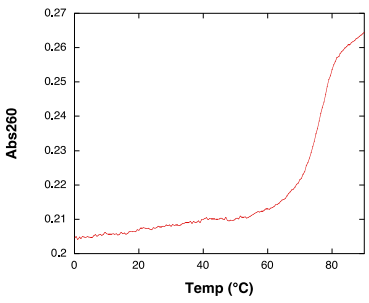
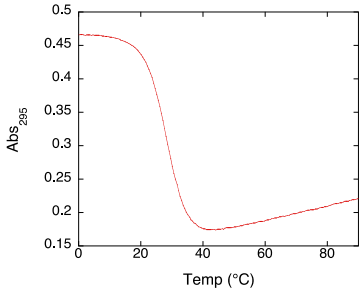
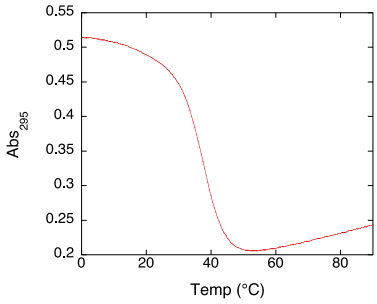

0.27
0.26
0.25
0.24
0.23
0.22
0.21
0.2
Abs260
0
20
40
60
80
Temp (°C)

Figure S7. UV melting curves measured at 295 nm at pH 6.0 in the absence of PEG for 10 µM oligonucleotides (A) $(C_3TA_2)_4$ in the absence of KCl, (B) Hif1a in the absence of KCl, (C) cILPR in the absence of KCl, (D) $(C_3TA_2)_4$ with 100 mM KCl, (E) Hif1a with 100 mM KCl, (F) cILPR with 100 mM KCl, (G) $(T_2AG_3)_4$ in the presence of 1 mM KCl, (H) $(T_2AG_3)_4$ in the presence of 10 mM KCl, (I) $(T_2AG_3)_4$ in the presence of 30 mM KCl, (J) $(T_2AG_3)_4$ in the presence of 50 mM KCl, (K) Q6 in the presence of 1 mM KCl, (L) ILPR in the absence of KCl, (M) H1 in the absence of KCl, (N) H2 in the absence of KCl, and (O) H3 in the absence of KCl. Oligonucleotide sequences are shown in Table 1. All solutions were buffered with 40 mM MES (pH 6.0) and 8 mM $MgCl_2$.

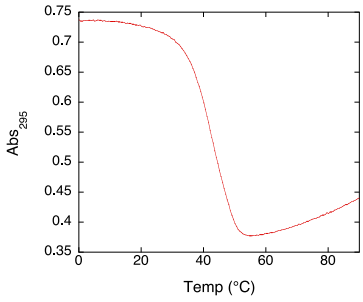
(A)


Abs295
Temp (°C)

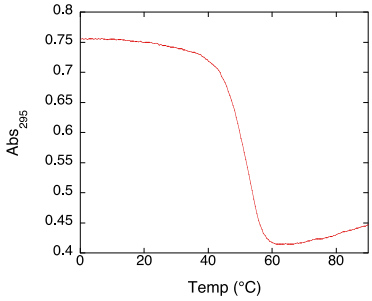
(B)


Abs295
Temp (°C)

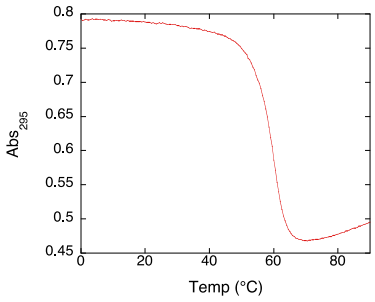
(C)


Abs295
Temp (°C)

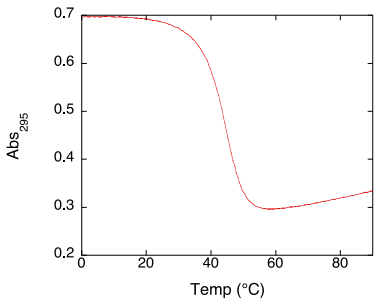
(D)


Abs295
Temp (°C)

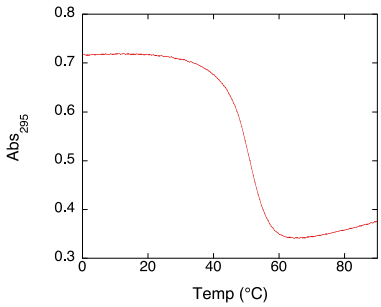
(E)


Abs295
Temp (°C)

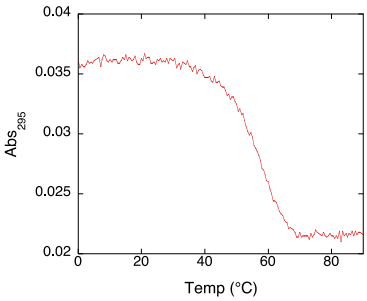
(F)


Abs295
Temp (°C)

(G)


Abs295
Temp (°C)

(H)


Abs295
Temp (°C)

(I)

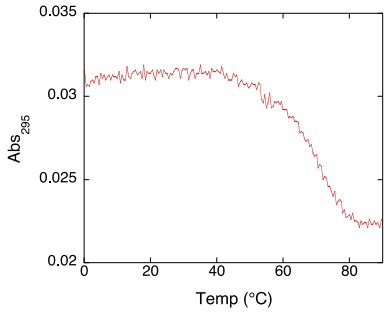


Figure S8. UV melting curves measured at 295 nm in the crowding conditions for 10 μM oligonucleotides (A) $(C_3TA_2)_4$ with 100 mM KCl and 20 wt% PEG200 at pH 6.0, (B) $(C_3TA_2)_4$ with 100 mM KCl and 20 wt% PEG1000 at pH 6.0, (C) Hif1a with 100 mM KCl and 20 wt% PEG200 at pH 6.0, (D) Hif1a with 100 mM KCl and 20 wt% PEG1000 at pH 6.0, (E) Hif1a with 100 mM KCl and 20 wt% PEG1000 at pH 7.0, (F) cILPR with 100 mM KCl and 20 wt% PEG200 at pH 6.0, (G) cILPR with 100 mM KCl and 20 wt% PEG1000 at pH 6.0, (H) $(T_2AG_3)_4$ in the presence of both 1 mM KCl and 20 wt% PEG200 at pH 6.0, and (I) $(T_2AG_3)_4$ in the presence of both 30 mM KCl and 20 wt% PEG200 at pH 6.0. Oligonucleotide sequences are shown in Table 1. All solutions were buffered with 40 mM MES (pH 6.0 or 7.0) and 8 mM $MgCl_2$.

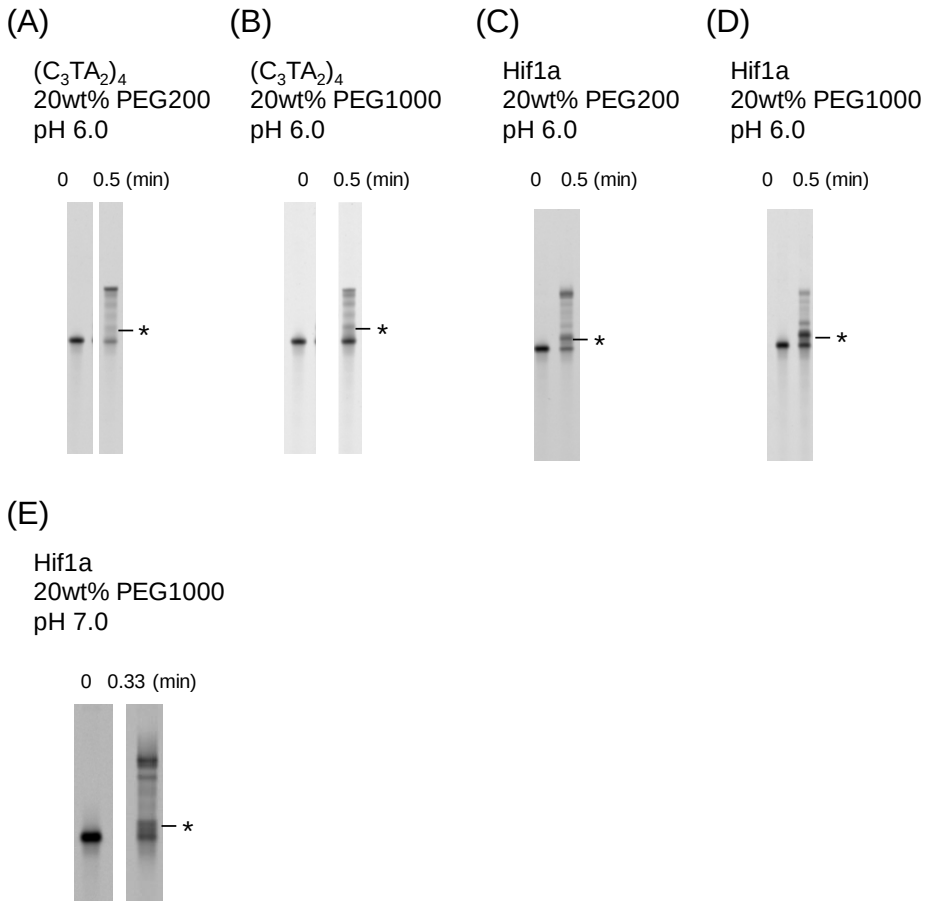


Figure S9. Effect of crowding condition on the replication. The reactions were sampled at 0 and 0.5 min in the case of pH 6.0 and at 0 and 0.33 min at pH 7.0, and run on denaturing PAGE. The gel images were recorded without staining, therefore fluorescently labeled primer and products are visualized. The fraction of replication completion was shown under the each gel image. (A) $(C_3TA_2)_4$ template with 20 wt% PEG200 at pH 6.0, (B) $(C_3TA_2)_4$ template with 20 wt% PEG1000 at pH 6.0, (C) Hif1a template with 20 wt% PEG200 at pH 6.0, (D) Hif1a template with 20 wt% PEG1000 at pH 6.0, and (E) Hif1a template with 20 wt PEG1000 pH 7.0. The asterisk on the right of each gel image indicates the migrated position of major stalled products. Reactions were carried out with 1 μM KF, 1 μM DNAs, and 25 μM dNTPs in the solution containing 40 mM MES (pH 6.0 or pH 7.0) and 8 mM $MgCl_2$ with 100 mM KCl at 37 °C.

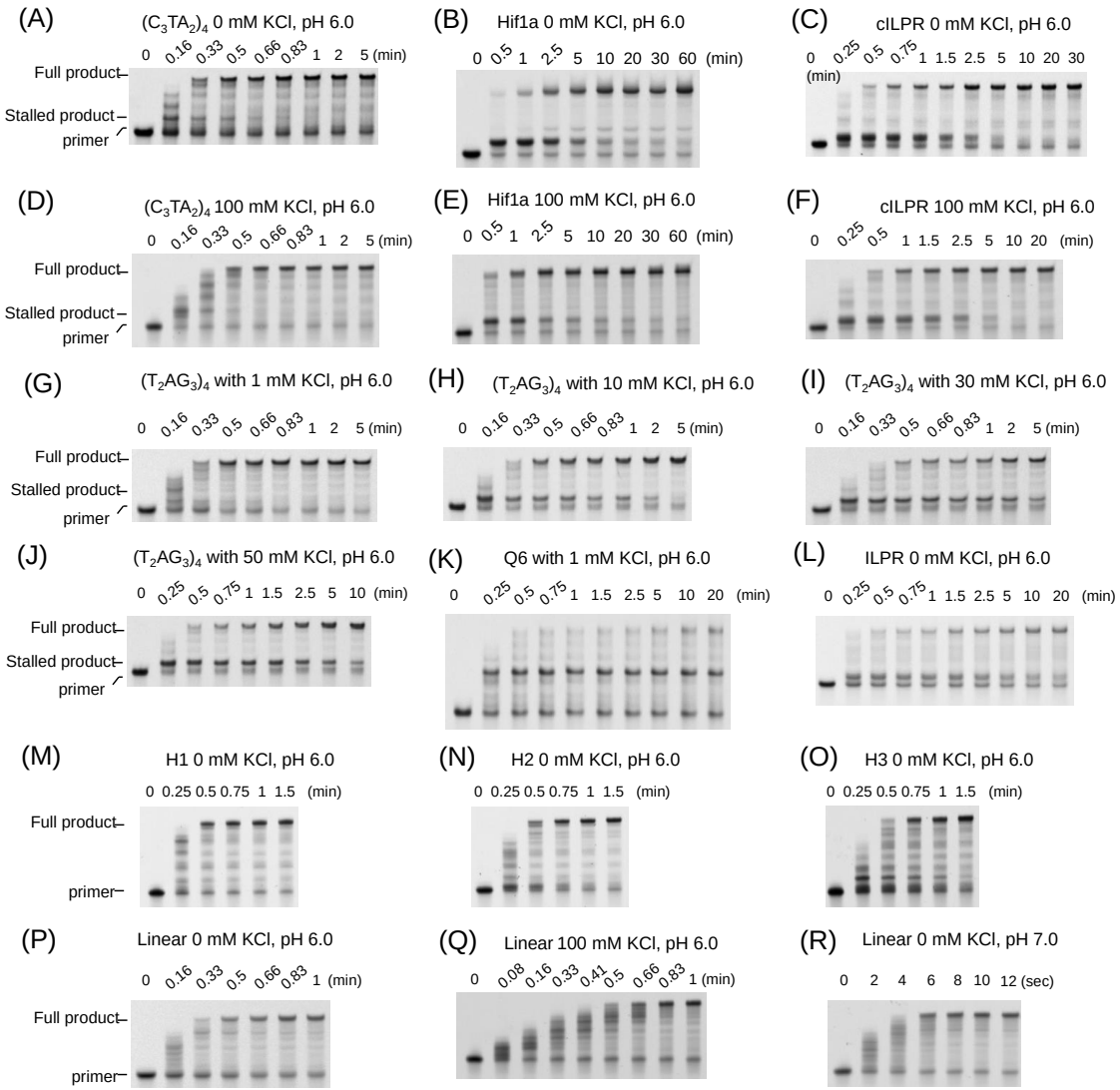


Figure S10. PAGE images of time courses of replication of (A) $(C_3TA_2)_4$ template in the absence of KCl at pH 6.0, (B) Hif1a template in the absence of KCl at pH 6.0, (C) cLIPR template in the absence of KCl at pH 6.0, (D) $(C_3TA_2)_4$ template in 100 mM KCl at pH 6.0, (E) Hif1a template in 100 mM KCl at pH 6.0, (F) cLIPR template in 100 mM KCl at pH 6.0, (G) $(T_2AG_3)_4$ template in 1 mM KCl at pH 6.0, (H) $(T_2AG_3)_4$ template in 10 mM KCl at pH 6.0, (I) $(T_2AG_3)_4$ template in 30 mM KCl at pH 6.0, (J) $(T_2AG_3)_4$ template in 50 mM KCl at pH 6.0, (K) Q6 template in 1 mM KCl at pH 6.0, (L) ILPR template in the absence of KCl at pH 6.0, (M) H1 in the absence of KCl at pH 6.0, (N) H2 in the absence of KCl at pH 6.0 (O) H3 in the absence of KCl at pH 6.0, (P) Linear in the absence of KCl at pH 6.0 (Q) Linear in 100 mM KCl at pH 6.0, and (R) Linear in 0 mM KCl at pH 7.0. All the reactions were carried in 40 mM MES (pH 6.0 or 7.0), 8 mM $MgCl_2$, 1 μM KF, 1 μM DNAs, and 250 μM dNTPs at 37 °C.

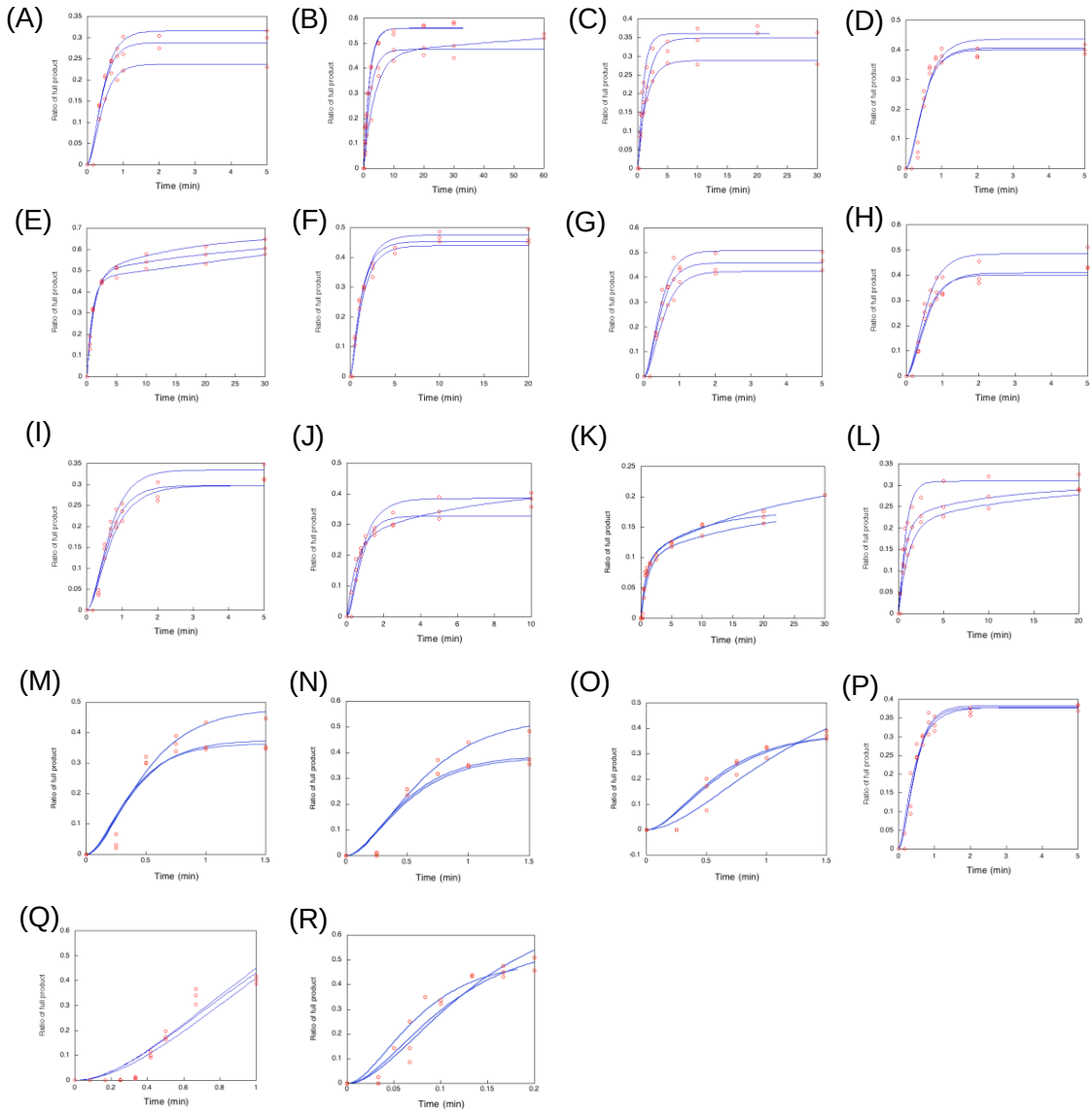


Figure S11. Analyses of full-length product formation in reactions in the absence of PEG with (A) $(C_3TA_2)_4$ template in the absence of KCl at pH 6.0, (B) Hif1a template in the absence of KCl at pH 6.0, (C) cLIPR template in the absence of KCl at pH 6.0, (D) $(C_3TA_2)_4$ template in 100 mM KCl at pH 6.0, (E) Hif1a template in 100 mM KCl at pH 6.0, (F) cLIPR template in 100 mM KCl at pH 6.0, (G) $(T_2AG_3)_4$ template in 1 mM KCl at pH 6.0, (H) $(T_2AG_3)_4$ template in 10 mM KCl at pH 6.0, (I) $(T_2AG_3)_4$ template in 30 mM KCl at pH 6.0, (J) $(T_2AG_3)_4$ template in 50 mM KCl at pH 6.0, (K) Q6 template in 1 mM KCl at pH 6.0, (L) ILPR template in the absence of KCl at pH 6.0, (M) H1 in the absence of KCl at pH 6.0, (N) H2 in the absence of KCl at pH 6.0, (O) H3 in the absence of KCl at pH

6.0, (P) Linear in the absence of KCl at pH 6.0, (Q) Linear in 100 mM KCl at pH 6.0, and (R) Linear in 0 mM KCl at pH 7.0. All solutions also contained 40 mM MES (pH 6.0) and 8 mM $MgCl_2$. All reactions were carried out in the presence of 1 μM KF and 1 μM DNAs at 37 °C.

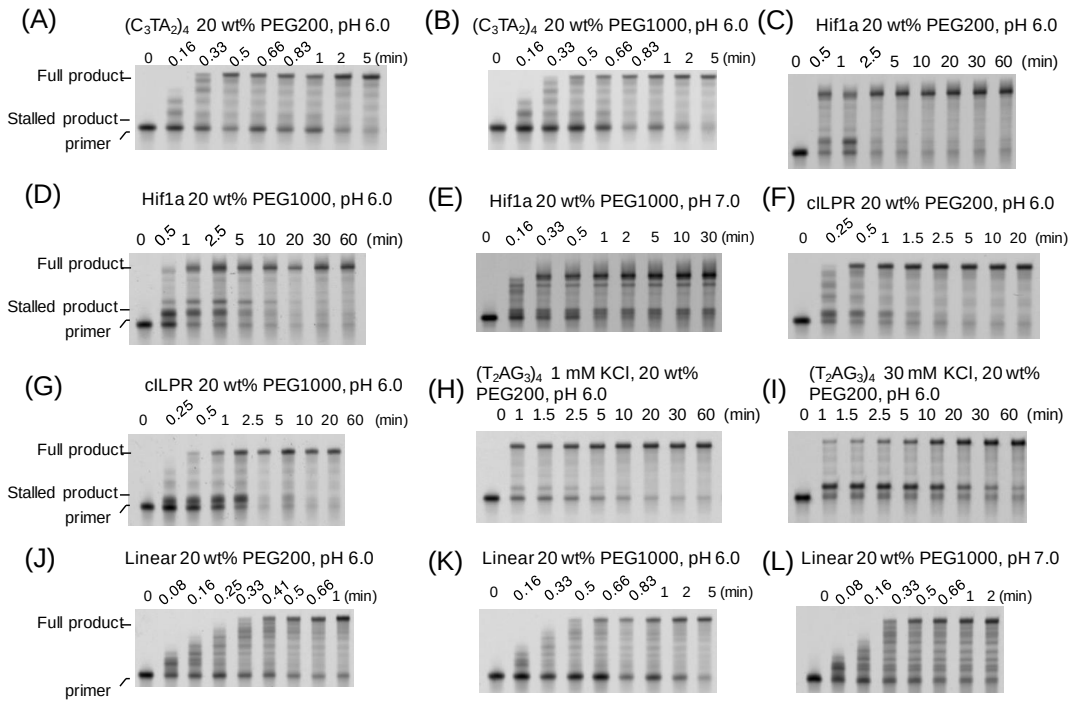


Figure S12. PAGE images of replication time courses in crowding conditions. (A) $(C_3TA_2)_4$ template in 20 wt% PEG200 at pH 6.0, (B) $(C_3TA_2)_4$ template in 20 wt% PEG1000 at pH 6.0, (C) Hif1a template in 20 wt% PEG200 at pH 6.0, (D) Hif1a template in 20 wt% PEG1000 at pH 6.0, (E) Hif1a template in 20 wt% PEG1000 at pH 7.0, (F) cLIPR template in 20 wt% PEG200 at pH 6.0, (G) cLIPR template in 20 wt% PEG1000 at pH 6.0, (H) $(T_2AG_3)_4$ template in 1 mM KCl and 20 wt% PEG200 at pH 6.0, (I) $(T_2AG_3)_4$ template in 30 mM KCl and 20 wt% PEG200 at pH 6.0, (J) Linear template in 20 wt% PEG200 at pH 6.0, (K) Linear template in 20 wt% PEG1000 at pH 6.0, and (L) Linear template in 20 wt% PEG1000 at pH 7.0. All the reactions were carried in 40 mM MES (pH 6.0 or 7.0), 8 mM $MgCl_2$, 100 mM KCl, 1 μM KF, 1 μM DNAs, and 250 μM dNTPs at 37 °C.

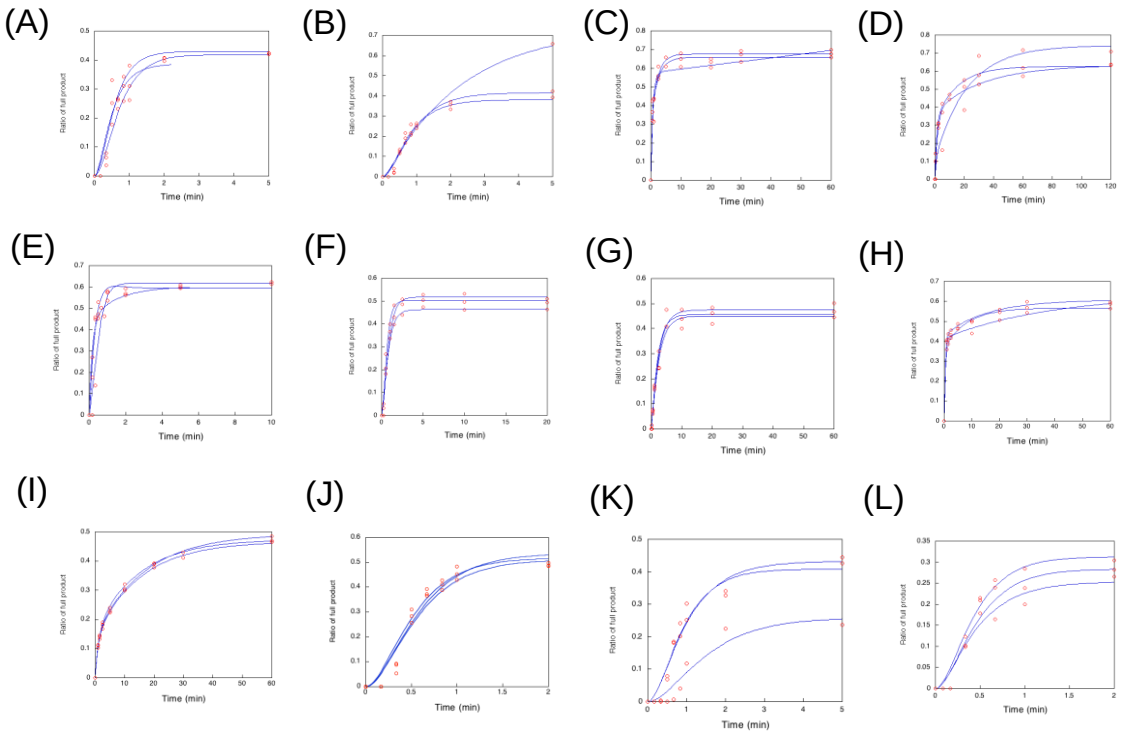


Figure S13. Analyses of full-length product formation in reactions in the presence of PEG with (A) $(C_3TA_2)_4$ template in 20 wt% PEG200 at pH 6.0, (B) $(C_3TA_2)_4$ template in 20 wt% PEG1000 at pH 6.0, (C) Hif1a template in 20wt% PEG200 at pH 6.0, (D) Hif1a template in 20 wt% PEG1000 at pH 6.0, (E) Hif1a template in 20 wt% PEG1000 at pH 7.0, (F) cLIPR template in 20 wt% PEG200 at pH 6.0, (G) cLIPR template in 20 wt% PEG1000 at pH 6.0, (H) $(T_2AG_3)_4$ template in 1 mM KCl and 20 wt% PEG200 at pH 6.0, (I) $(T_2AG_3)_4$ template in 30 mM KCl and 20 wt% PEG200 at pH 6.0, (J) Linear template in 20 wt% PEG200 at pH 6.0, (K) Linear template in 20 wt% PEG1000 at pH 6.0, and (L) Linear template in 20 wt% PEG1000 at pH 7.0. All solutions also contained 40 mM MES (pH 6.0 or 7.0), 8 mM $MgCl_2$, 100 mM KCl, and 20 wt% of PEG200 or PEG1000. All reactions were carried out in the presence of 1 μM KF and 1 μM DNAs at 37 °C.

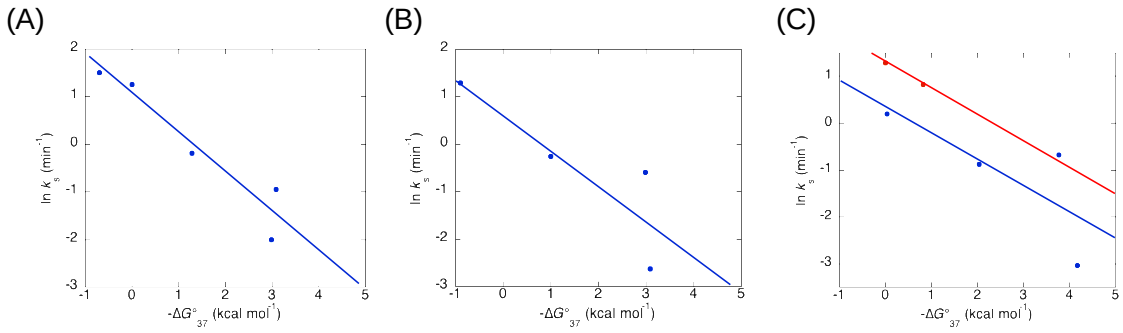


Figure S14. lnk_s and -ΔG°$_{37}$ plot for replication of i-motif-forming DNA at pH 6.0 (A) in the absence of KCl and PEG at pH 6.0, (B) in the presence of and 100 mM KCl at pH 6.0, and (C) in the presence of 20 wt% PEG1000 and 100 mM KCl at pH 6.0 (blue) and 7.0 (red). All the reactions were carried out with 40 mM MES (pH 6.0 or 7.0), 8 mM $MgCl_2$, 1 μM KF, and 250 μM dNTPs, together with the designated KCl and PEG conditions.

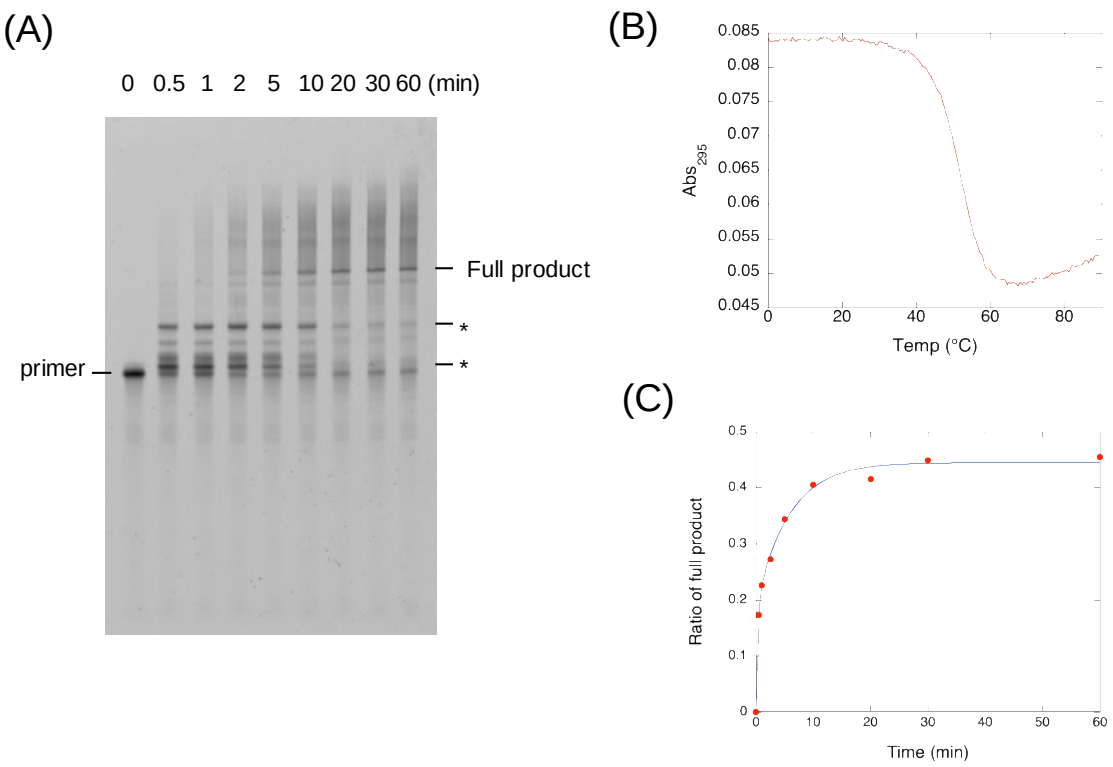


Figure S15. Effect of i-motif topology on replication. (A) PAGE images of products of replication of cBcl2 template DNA. The asterisk indicates the major stall products. All the reactions were carried out in 40 mM MES (pH 6.0), 8 mM $MgCl_2$, 1 µM KF, 1 µM DNA template and primer, and 250 µM dNTPs at 37 °C. (B) UV melting curves measured at 295 nm at pH 6.0 for cBcl2. The concentration of DNAs was 10 µM. The solution was buffered with 40 mM MES (pH 6.0) and 8 mM $MgCl_2$. (C) Analyses of full-length product formation based on experiment shown in (A).

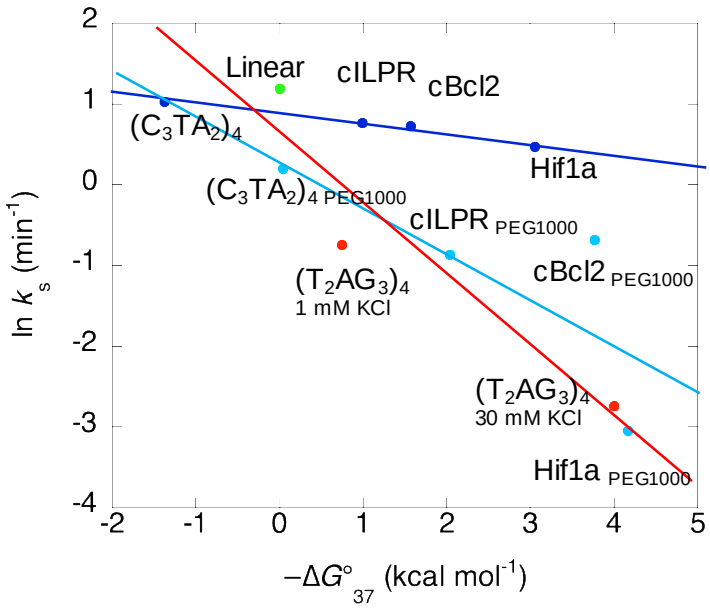


Figure S16. Plot for the logarithms of rate constants for reactions versus -$\Delta G°_{37}$ values to produce full-length replication product from i-motif-forming templates (blue: in the presence of 20 wt% PEG200, light blue: in the presence of 20 wt% PEG1000), G4-forming templates (red), and Linear template (green). The blue and light-blue lines indicate the linear correlation between lnk_s and -$\Delta G°_{37}$ of i-motif-forming sequences in the presence of 20 wt% PEG200 and 20 wt% PEG1000, respectively. The red line indicates the linear correlation between lnk_s and -$\Delta G°_{37}$ of $(T_2AG_3)_4$ G4. In the presence of 20 wt% PEG1000, the addition of 100 mM KCl is required to avoid KF denaturation. In that condition, the stability of $(T_2AG_3)_4$ was too high to determine. All the reactions were performed with 40 mM MES (pH 6.0), 8 mM $MgCl_2$, 100 mM KCl, 20 wt% PEG200, 1 μM KF, and 250 μM dNTPs.